GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 09:22:47 ; Search time 100.93 Seconds
(without alignments)
925.874 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	98	89.9	2379	21	AAA96354	cDNA encoding a no
	2	97	89.0	144	19	AAV21239	Homo sapiens clone
	3	97	89.0	144	20	AAX77331	Human secreted pro
c	4	97	89.0	170	22	AAH70047	Human cervical can
	5	97	89.0	378	22	AAS01025	Sugarcane plant ge
c	6	97	89.0	396	22	AAF94819	Human ovarian canc
c	7	97	89.0	597	22	AAH71472	Human cervical can
c	8	97	89.0	621	22	AAH71551	Human cervical can
	9	97	89.0	801	21	AAA64638	Partial sequence M
	10	97	89.0	976	22	AAS01021	Sugarcane plant ge
	11	97	89.0	1034	21	AAZ52527	Human secreted pro
	12	97	89.0	1117	21	AAA59475	Nucleotide sequenc
	13	97	89.0	1164	22	AAF98699	Human ovarian canc
	14	97	89.0	1376	20	AAX04325	Human secreted pro
	15	97	89.0	1495	21	AAA27985	Corn CCR4 transcri
	16	97	89.0	1763	21	AAA27151	Human p52 cDNA. H
	17	97	89.0	1812	21	AAA16649	Human secreted pro
	18	97	89.0	1936	21	AAZ43798	Human fetal brain
	19	97	89.0	1962	22	AAF30434	Corn apoptosis ind
	20	97	89.0	2338	21	AAF18071	Lung cancer associ
	21	97	89.0	3030	21	AAC61886	cDNA encoding a hu
	22	97	89.0	3501	22	AAS21262	Human cDNA sequenc
	23	97	89.0	4064	21	AAA07587	Mouse piwi gene, d
	24	97	89.0	4640	21	AAC78483	Human PRO708 (UNQ3
	25	97	89.0	4650	20	AAZ33987	Human PRO708 nucle
	26	96	88.1	158	22	AAS23328	Human prostate can
	27	96	88.1	225	21	AAC98734	Human colon cancer
c	28	96	88.1	310	22	AAH71505	Human cervical can
c	29	96	88.1	396	22	AAF94842	Human ovarian canc
	30	96	88.1	451	22	AAI57612	Human colorectal c
	31	96	88.1	464	22	AAH33241	Human colon cancer
	32	96	88.1	569	21	AAZ97052	Human secreted pro
	33	96	88.1	612	21	AAC99862	Human secreted pro
	34	96	88.1	772	19	AAV58363	Coding sequence fo
	35	96	88.1	936	20	AAZ06226	Human secreted pro
	36	96	88.1	938	22	AAD07722	Human secreted pro
	37	96	88.1	1024	21	AAZ97189	Human prostate can
	38	96	88.1	1048	21	AAA26373	Human secreted pro
	39	96	88.1	1051	21	AAZ91918	Murine mahogany pr
	40	96	88.1	1126	21	AAC59409	Human secreted pro
	41	96	88.1	1145	22	AAS02426	Human secreted pro
	42	96	88.1	1149	21	AAC81030	Human secreted pro
	43	96	88.1	1198	22	AAF91902	Human secreted pro
	44	96	88.1	1254	21	AAC99826	Human secreted pro
	45	96	88.1	1474	18	AAT90174	Oil seed rape cyst

ALIGNMENTS

RESULT 1
AAA96354
ID AAA96354 standard; cDNA; 2379 BP.
XX
AC AAA96354;
XX
DT 08-FEB-2001 (first entry)
XX
DE cDNA encoding a novel polypeptide designated PRO6030.
XX
KW Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122;
KW PRO1889; PRO1890; PRO1887; PRO1785; PRO4353; PRO4357; PRO4405; PRO4356;
KW PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030;
KW PRO4424; PRO4422; PRO4430; PRO4499; tumour; obesity; diabetes;
KW insulinemia; kidney disorder; Bergers disease; nephropathy;
KW Schonlein-Henoch purpura; celiac disease; dermatitis herpetiformis;
KW Crohns disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 60..1028
FT /*tag= a
XX
PN WO200056889-A2.
XX
PD 28-SEP-2000.
XX
PF 01-MAR-2000; 2000WO-US05601.
XX
PR 23-MAR-1999; 99US-0125774.
PR 23-MAR-1999; 99US-0125778.
PR 24-MAR-1999; 99US-0125826.
PR 31-MAR-1999; 99US-0127035.

PR 05-APR-1999; 99US-0127706.
PR 21-APR-1999; 99US-0130359.
PR 27-APR-1999; 99US-0131270.
PR 27-APR-1999; 99US-0131272.
PR 27-APR-1999; 99US-0131291.
PR 04-MAY-1999; 99US-0132371.
PR 04-MAY-1999; 99US-0132379.
PR 04-MAY-1999; 99US-0132383.
PR 25-MAY-1999; 99US-0135750.
PR 08-JUN-1999; 99US-0138166.
PR 20-JUL-1999; 99US-0144791.
PR 03-AUG-1999; 99US-0146970.
PR 09-DEC-1999; 99US-0170262.
XX
PA (GETH) GENENTECH INC.
XX
PI Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
PI Stewart TA, Watanabe CK, Wood WI, Zhang Z;
XX
DR WPI; 2000-628263/60.
DR P-PSDB; AAB18927.
XX
PT Novel secreted and transmembrane polypeptides useful for diagnosing
PT tumour in a mammal, for identifying agonists and antagonists of the
PT polypeptide and for therapeutic use -
XX
PS Claim 2; Fig 37; 222pp; English.
XX
CC The present sequence encodes a secreted or transmembrane polypeptide.
CC The specification describes polypeptides designated PRO1484, PRO4334,
CC PRO1122, PRO1889, PRO1890, PRO1887, PRO1785, PRO4353, PRO4357, PRO4405,
CC PRO4356, PRO4352, PRO4380, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990,
CC PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is
CC useful for diagnosing tumour in a mammal. The polypeptides, their
CC agonists and antagonists are useful treating a condition associated with
CC expression or activity of the polypeptide. Conditions treated include
CC obesity, diabetes or hyper-or hypo-insulinemia. The polypeptides are
CC capable of inducing proliferation of mammalian kidney mesangial cells
CC and are therefore useful for treating kidney disorders associated with
CC decreased mesangial cell function such as Bergers disease or other
CC nephropathies associated with Schonlein-Henoch purpura, celiac disease,
CC dermatitis herpetiformis or Crohns disease. The nucleic acids may be used
CC to generate transgenic animals for use in development and screening of
CC therapeutically useful reagents and also for chromosome identification
CC and tissue typing.
XX
SQ Sequence 2379 BP; 574 A; 695 C; 604 G; 506 T; 0 other;

Query Match 89.9%; Score 98; DB 21; Length 2379;
Best Local Similarity 100.0%; Pred. No. 8e-17;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 2277 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2336

Qy 70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 2337 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2374

RESULT 2
AAV21239
ID AAV21239 standard; cDNA; 144 BP.
XX
AC AAV21239;
XX
DT 03-AUG-1998 (first entry)
XX
DE Homo sapiens clone H698_3 3' end.
XX
KW secreted protein; human; PBMC; peripheral blood mononuclear cells; ds.
XX
OS Homo sapiens.
XX
PN WO9807859-A2.
XX
PD 26-FEB-1998.
XX
PF 22-AUG-1997; 97WO-US14874.
XX
PR 23-AUG-1996; 96US-0702344.
XX
PA (GEMY) Genetics Inst Inc.
XX
PI Merberg F, McCoy JM, Lavallie ER, Racie LA, Treacy M, Spaulding V;
PI Jacobs K;
XX
DR WPI; 1998-169163/15.
XX
PT New nucleic acid encoding secreted proteins from human cells -
PT useful e.g. as immuno-modulators, antitumour agents, promoters of
PT tissue growth, haemostatic and thrombolytic agents etc.
XX
PS Claim 33; Page 62; 79pp; English.
XX
CC The sequence is that of the 3' end of clone H698_3 (including
CC the polyA tail) which was isolated from a human adult PBMC cDNA
CC library using methods selective for cDNAs that encode secreted
CC proteins.
XX
SQ Sequence 144 BP; 132 A; 0 C; 0 G; 8 T; 4 other;

Query Match 89.0%; Score 97; DB 19; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 20 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 79

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 80 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 116

RESULT 3
AAX77331
ID AAX77331 standard; DNA; 144 BP.
XX
AC AAX77331;
XX
DT 04-AUG-1999 (first entry)
XX
DE Human secreted protein encoding DNA (clone H698-3) 3' portion.
XX
KW Secreted protein; retina; brain; blood; testis; nutritional activity;
KW cell proliferation; differentiation; immune stimulation; human;
KW hematopoiesis regulation; tissue growth; chemotactic; thrombolytic;
KW anti-inflammatory; tumour invasion suppression; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO9926973-A1.
XX
PD 03-JUN-1999.
XX
PF 20-NOV-1998; 98WO-US24944.
XX
PR 19-NOV-1998; 98US-0196027.
PR 21-NOV-1997; 97US-0976112.
XX
PA (GEMY) GENETICS INST INC.
XX

PI Collins-Racie LA, Evans C, Jacobs K, Lavallie ER;
PI McCoy JM, Merberg D, Spaulding V, Treacy M;
XX
DR WPI; 1999-357814/30.
XX
PT New polynucleotides encoding secreted human proteins
XX
PS Disclosure; Page 97; 106pp; English.
XX
CC The invention relates to human secreted proteins (AAY21588-597) encoded
CC by polynucleotides obtained from adult retina, adult brain, adult blood
CC or adult testes cDNA libraries. The nucleic acid sequences encoding the
CC secreted proteins correspond to clones AY421-2, BV278-2, C544-1,
CC CC332-33, CC365-40, CG68-4, D329-1, H698-2 and H963-20, (all clones are
CC deposited as ATCC 98145). The PNs and proteins are predicted to have
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and animals,
CC although no supporting data is given. Suggested activities include
CC nutritional activity, cytokine and cell proliferation/differentiation
CC activity, immune stimulating (e.g. as vaccines) or suppressing activity,
CC hematopoiesis regulating activity, tissue growth activity, activin/
CC inhibin activity, chemotactic/chemokinetic activity, hemostatic and
CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumor invasion suppressor activity and tumor
CC inhibition activity. The PNs are also stated to be useful for gene
CC therapy.
XX
SQ Sequence 144 BP; 132 A; 0 C; 0 G; 8 T; 4 other;

Query Match 89.0%; Score 97; DB 20; Length 144;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 20 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 79

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 80 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 116

RESULT 4
AAH70047/c
ID AAH70047 standard; cDNA; 170 BP.
XX
AC AAH70047;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 1321.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US33312.
XX
PR 08-DEC-1999; 99US-0169681.
PR 21-DEC-1999; 99US-0171350.
PR 14-MAR-2000; 2000US-0189315.
PR 12-MAY-2000; 2000US-0203791.
PR 09-JUN-2000; 2000US-0210600.
PR 21-JUL-2000; 2000US-0220114.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX
PT New isolated nucleic acid for diagnosing and treating cervical cancer
PT and for assessing and detecting compounds for treating the cancer -
XX
PS Claim 1; Page 309; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy.
XX
SQ Sequence 170 BP; 11 A; 11 C; 10 G; 137 T; 1 other;

Query Match 89.0%; Score 97; DB 22; Length 170;
Best Local Similarity 100.0%; Pred. No. 4e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 163 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 104

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 103 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 67

RESULT 5
AAS01025
ID AAS01025 standard; cDNA; 378 BP.
XX
AC AAS01025;
XX
DT 29-MAY-2001 (first entry)
XX
DE Sugarcane plant gene promoter cDNA isolated from clone c32A.
XX
KW Sugarcane promoter region; monocotyledonous plant; stem tissue;
KW insecticide; herbicide; disease resistance; improved food content;
KW beta-glucuronidase; GUS; starch biosynthesis; fatty acid biosynthesis;
KW ADP-glucose pyrophosphorylase; sucrose metabolism; clone c32A; ss.
XX
OS Saccharum sp.
XX
FH Key Location/Qualifiers
FT CDS 2..46
FT /*tag= a
FT /partial
FT /note= "Amino acid sequence deduced from the
FT longest and conserved ORF. This sequence lacks
FT a start codon"
XX
PN WO200118211-A1.
XX
PD 15-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-AU01033.
XX
PR 02-SEP-1999; 99AU-0002625.
XX
PA (UYQU) UNIV QUEENSLAND.
XX
PI Potier B, Birch RG;
XX
DR WPI; 2001-218560/22.
DR P-PSDB; AAU00452.
XX

```
PT   New sugarcane plant promoters for directing expression of heterologous
PT   nucleic acids in a constitutive or tissue-specific manner in
PT   monocotyledonous plants -
XX
PS   Claim 1; Fig 16; 107pp; English.
XX
CC   The present sequence for sugarcane plant promoter cDNA isolated from
CC   clone c32A is 1 of 11 promoter regions of a transcribable DNA sequence
CC   isolated from various sugarcane cDNA clones (AAS01021-AAS01031). Also
CC   described are 4 promoter regions of specific transcribed DNA sequences
CC   (AAS01032-AAS01035). The nucleic acids are useful for producing
CC   transgenic plants, having an altered phenotype and for driving expression
CC   of a foreign or endogenous DNA sequence, which encode agronomic
CC   properties including insecticide, herbicide, disease resistance,
CC   stress tolerance and improved food content, or increased yields. The
CC   foreign or endogenous DNA sequence may comprise a region transcribed
CC   into an antisense RNA or ribozyme that modulates the expression of a
CC   corresponding target gene, or it may encode beta-glucuronidase (GUS),
CC   luciferase, neomycin phosphotransferase, a product conferring herbicide
CC   tolerance, a product affecting starch biosynthesis or modification,
CC   ADP-glucose pyrophosphorylase, a product involved in fatty acid
CC   biosynthesis, a product conferring insect resistance, a product altering
CC   sucrose metabolism or a gene encoding valuable pharmaceuticals, e.g.
CC   antibiotics, secondary metabolites or vaccines. The promoters are capable
CC   of directing high level expression in many or all cells of a plant,
CC   preferentially in stem or meristem tissue of monocotyledonous plants.
XX
SQ   Sequence 378 BP; 154 A; 49 C; 78 G; 97 T; 0 other;


  Query Match             89.0%;  Score 97;  DB 22;  Length 378;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     277 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 336

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db     337 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 373


RESULT   6
AAF94819/c
ID   AAF94819 standard; cDNA; 396 BP.
XX
AC   AAF94819;
XX
DT   23-MAY-2001  (first entry)
XX
DE   Human ovarian cancer associated coding sequence SEQ ID NO: 10.
XX
KW   Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200118046-A2.
XX
PD   15-MAR-2001.
XX
PF   08-SEP-2000; 2000WO-US24827.
XX
PR   10-SEP-1999;   99US-0394374.
PR   01-MAY-2000; 2000US-0561778.
PR   15-AUG-2000; 2000US-0640173.
PR   07-SEP-2000; 2000US-0656668.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Xu J,  Stolk JA;
XX
DR   WPI; 2001-211395/21.
XX
PT   Isolated polypeptides associated with ovarian carcinomas, and the
PT   nucleic acids that encode them, useful for the prevention diagnosis and
PT   treatment of ovarian cancers -
XX
PS   Claim 5; Page 119; 189pp; English.
XX
CC   The present invention provides a number of coding sequences and proteins,
CC   the over-expression of which is associated with ovarian carcinoma/cancer.
CC   These can be used in the diagnosis, treatment and prevention of ovarian
CC   cancer, optionally by gene therapy or in the form of a vaccine. The
CC   present sequence is an example of one of these sequences.
XX
SQ   Sequence 396 BP; 85 A; 18 C; 55 G; 153 T; 85 other;


  Query Match             89.0%;  Score 97;  DB 22;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 48

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db      47 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11


RESULT   7
AAH71472/c
ID   AAH71472 standard; cDNA; 597 BP.
XX
AC   AAH71472;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 2746.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US33312.
XX
PR   08-DEC-1999;   99US-0169681.
PR   21-DEC-1999;   99US-0171350.
PR   14-MAR-2000; 2000US-0189315.
PR   12-MAY-2000; 2000US-0203791.
PR   09-JUN-2000; 2000US-0210600.
PR   21-JUL-2000; 2000US-0220114.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer
PT   and for assessing and detecting compounds for treating the cancer -
XX
PS   Claim 1; Page 564; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
```

```
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy.
XX
SQ   Sequence 597 BP; 259 A; 54 C; 39 G; 244 T; 1 other;


  Query Match             89.0%;  Score 97;  DB 22;  Length 597;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     141 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 82

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||
Db      81 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 45


RESULT   8
AAH71551/c
ID   AAH71551 standard; cDNA; 621 BP.
XX
AC   AAH71551;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 2825.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US33312.
XX
PR   08-DEC-1999;   99US-0169681.
PR   21-DEC-1999;   99US-0171350.
PR   14-MAR-2000; 2000US-0189315.
PR   12-MAY-2000; 2000US-0203791.
PR   09-JUN-2000; 2000US-0210600.
PR   21-JUL-2000; 2000US-0220114.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer
PT   and for assessing and detecting compounds for treating the cancer -
XX
PS   Claim 1; Page 579; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy.
XX
SQ   Sequence 621 BP; 273 A; 63 C; 47 G; 238 T; 0 other;


  Query Match             89.0%;  Score 97;  DB 22;  Length 621;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 76

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||
Db      75 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 39


RESULT   9
AAA64638
ID   AAA64638 standard; DNA; 801 BP.
XX
AC   AAA64638;
XX
DT   02-JAN-2001  (first entry)
XX
DE   Partial sequence MEL3 associated with melanoma and thyroid tumors.
XX
KW   Melanoma; thyroid tumour; MEL3; MEL7; Thy5; Thy6; Thy11; Thy14; Thy15;
KW   rectal cancer; lung cancer; breast cancer; colon cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200050595-A2.
XX
PD   31-AUG-2000.
XX
PF   25-FEB-2000; 2000WO-US04929.
XX
PR   25-FEB-1999;   99US-0257417.
XX
PA   (GOUT/) GOUT I.
PA   (RODN/) RODNIN N.
PA   (FILO/) FILONENKO V.
PA   (MATS/) MATSUKA G.
PA   (SCAN/) SCANLAN M.
PA   (OLDL/) OLD L.
PA   (BILY/) BILYNSKY B.
XX
PI   Gout I,  Rodnin N,  Filonenko V,  Matsuka G,  Scanlan M,  Old L;
PI   Bilynsky B;
XX
DR   WPI; 2000-572092/53.
XX
PT   Novel isolated nucleic acid molecules for diagnosing and treating
PT   melanoma, thyroid tumors, rectal, lung, breast and colon cancers  -
XX
PS   Claim 1; Page 62; 94pp; English.
XX
CC   The specification describes polynucleotides which are associated with
CC   melanoma and thyroid tumors. Specifically, these are MEL3, MEL7, Thy5,
CC   Thy6, Thy11, Thy14, and Thy15. The polynucleotides are useful for
CC   diagnosing and treating a patient with melanoma, thyroid tumour,
CC   rectal cancer, lung cancer, breast cancer or colon cancer. The
CC   present sequence represents a polynucleotide of the invention.
XX
SQ   Sequence 801 BP; 178 A; 198 C; 162 G; 165 T; 98 other;


  Query Match             89.0%;  Score 97;  DB 21;  Length 801;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     648 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 707

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||
Db     708 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 744
```

```
RESULT 10
AAS01021
ID   AAS01021 standard; cDNA; 976 BP.
XX
AC   AAS01021;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Sugarcane plant gene promoter cDNA isolated from clone c67.
XX
KW   Sugarcane promoter region; monocotyledonous plant; stem tissue;
KW   insecticide; herbicide; disease resistance; improved food content;
KW   beta-glucuronidase; GUS; starch biosynthesis; fatty acid biosynthesis;
KW   ADP-glucose pyrophosphorylase; sucrose metabolism; clone c67; ss.
XX
OS   Saccharum sp.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..32
FT                   /*tag=  a
FT   CDS             33..596
FT                   /*tag=  b
FT                   /note= "Amino acid sequence deduced from the
FT                        longest and conserved ORF"
FT   3'UTR           597..581
FT                   /*tag=  c
FT   polyA_site      882
FT                   /*tag=  d
XX
PN   WO200118211-A1.
XX
PD   15-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-AU01033.
XX
PR   02-SEP-1999;   99AU-0002625.
XX
PA   (UYQU ) UNIV QUEENSLAND.
XX
PI   Potier B,  Birch RG;
XX
DR   WPI; 2001-218560/22.
DR   P-PSDB; AAU00450.
XX
PT   New sugarcane plant promoters for directing expression of heterologous
PT   nucleic acids in a constitutive or tissue-specific manner in
PT   monocotyledonous plants  -
XX
PS   Claim 1; Fig 14; 107pp; English.
XX
CC   The present sequence for sugarcane plant promoter cDNA isolated from
CC   clone c67 is 1 of 11 promoter regions of a transcribable DNA sequence
CC   isolated from various sugarcane cDNA clones (AAS01021-AAS01031). Also
CC   described are 4 promoter regions of specific transcribed DNA sequences
CC   (AAS01032-AAS01035). The nucleic acids are useful for producing
CC   transgenic plants, having an altered phenotype and for driving expression
CC   of a foreign or endogenous DNA sequence, which encode agronomic
CC   properties including insecticide, herbicide, disease resistance,
CC   stress tolerance and improved food content, or increased yields. The
CC   foreign or endogenous DNA sequence may comprise a region transcribed
CC   into an antisense RNA or ribozyme that modulates the expression of a
CC   corresponding target gene, or it may encode beta-glucuronidase (GUS),
CC   luciferase, neomycin phosphotransferase, a product conferring herbicide
CC   tolerance, a product affecting starch biosynthesis or modification,
CC   ADP-glucose pyrophosphorylase, a product involved in fatty acid
CC   biosynthesis, a product conferring insect resistance, a product altering
CC   sucrose metabolism or a gene encoding valuable pharmaceuticals, e.g.
CC   antibiotics, secondary metabolites or vaccines. The promoters are capable
CC   of directing high level expression in many or all cells of a plant,
CC   preferentially in stem or meristem tissue of monocotyledonous plants.
XX
SQ   Sequence 976 BP; 320 A; 213 C; 229 G; 214 T; 0 other;

  Query Match             89.0%;  Score 97;  DB 22;  Length 976;
  Best Local Similarity   100.0%;  Pred. No. 2e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     879 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 938

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db     939 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 975


RESULT 11
AAZ52527
ID   AAZ52527 standard; cDNA; 1034 BP.
XX
AC   AAZ52527;
XX
DT   29-FEB-2000  (first entry)
XX
DE   Human secreted protein clone ya66_1 nucleotide sequence SEQ ID NO:105.
XX
KW   Human; secreted protein; immunostimulatory; haemostatic; cytokine;
KW   proliferative; differentiative; chemotactic; chemokinetic; vaccine;
KW   thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
KW   gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO9958642-A2.
XX
PD   18-NOV-1999.
XX
PF   14-MAY-1999;   99WO-US10843.
XX
PR   14-MAY-1998;   98US-0085472.
PR   17-AUG-1998;   98US-0096824.
PR   11-SEP-1998;   98US-0099843.
PR   11-SEP-1998;   98US-0099950.
PR   15-SEP-1998;   98US-0100424.
PR   29-SEP-1998;   98US-0102329.
PR   09-OCT-1998;   98US-0103615.
PR   11-DEC-1998;   98US-0111799.
PR   14-DEC-1998;   98US-0112159.
PR   31-DEC-1998;   98US-0114415.
PR   10-FEB-1999;   99US-0248059.
PR   06-APR-1999;   99US-0287150.
PR   13-MAY-1999;   99US-0311021.
XX
PA   (GEMY ) GENETICS INST INC.
XX
PI   Wong GG,  Clark HF,  Fechtel K,  Agostino MJ;
XX
DR   WPI; 2000-053095/04.
DR   P-PSDB; AAY73442.
XX
PT   Novel polynucleotides and proteins having biological activities which
PT   make them suitable for treating, preventing or ameliorating medical
PT   conditions in humans or animals -
XX
PS   Claim 114; Page 653; 730pp; English.
XX
CC   The present invention describes human secreted proteins encoded by
CC   polynucleotides obtained from adult testes, foetal brain, adult brain,
CC   brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
CC   cDNA libraries. The polynucleotides and proteins are predicted to have
CC   biological activities which would make them suitable for treating,
CC   preventing or ameliorating medical conditions in humans and animals.
CC   Suggested activities include nutritional activity, cytokine and cell
CC   proliferation/differentiation activity, immune stimulating (e.g. as
CC   vaccines) or suppressing activity, haematopoiesis regulating activity,
```

CC tissue growth activity, activin/inhibin activity, chemotactic/
CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
CC suppressor activity, and tumour inhibition activity. The polynucleotides
CC are also stated to be useful for gene therapy. Therapeutic compositions
CC are also presently valuable for veterinary applications. AAZ52475 to
CC AAZ52581 encode human secreted proteins, and AAY73390 to AAY73500
CC represent human secreted proteins, given in the present invention.
XX
SQ Sequence 1034 BP; 305 A; 228 C; 222 G; 276 T; 3 other;

Query Match 89.0%; Score 97; DB 21; Length 1034;
Best Local Similarity 100.0%; Pred. No. 1.9e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 922 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 981

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 982 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1018

RESULT 12
AAA59475
ID AAA59475 standard; DNA; 1117 BP.
XX
AC AAA59475;
XX
DT 14-NOV-2000 (first entry)
XX
DE Nucleotide sequence of a soybean type III glutathione-S-transferase.
XX
KW Soybean; glutathione-S-transferase; GST; detoxification;
KW xenobiotic compound; herbicide-tolerance; transgenic plant;
KW herbicide synergist; ss.
XX
OS Glycine max.
XX
FH Key Location/Qualifiers
FT CDS 86..760
FT /*tag= a
FT /product= "glutathione-S-transferase"
XX
PN WO200047728-A2.
XX
PD 17-AUG-2000.
XX
PF 10-FEB-2000; 2000WO-US03347.
XX
PR 10-FEB-1999; 99US-0247373.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI McGonigle B, O'Keefe DP;
XX
DR WPI; 2000-549144/50.
DR P-PSDB; AAB07840.
XX
PT Soybean glutathione-S-transferase polypeptides and polynucleotides used
PT to produce herbicide tolerant transgenic plants and to screen for
PT inhibitors or substrates of the enzyme -
XX
PS Claim 24; Page 71; 84pp; English.
XX
CC The present sequence encodes a soybean glutathione-S-transferase
CC (GST) enzyme. The enzyme is involved in the detoxification of
CC xenobiotic compounds in plants and seeds. The GST polynucleotides
CC and polypeptides are used for the production of herbicide-tolerant
CC transgenic plants, and for the development of screening assays to
CC identify GST inhibitors and substrates, which can be used as
CC herbicide synergists. GST Gene specific probes can be used in gene
CC identification methods. The recombinant GST enzymes can be used to
CC produce enzyme specific antibodies which are used to detect the
CC enzymes in situ in cells or in vitro in cell extracts.
XX
SQ Sequence 1117 BP; 444 A; 153 C; 221 G; 296 T; 3 other;

Query Match 89.0%; Score 97; DB 21; Length 1117;
Best Local Similarity 100.0%; Pred. No. 1.9e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 994 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1053

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 1054 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1090

RESULT 13
AAF98699
ID AAF98699 standard; DNA; 1164 BP.
XX
AC AAF98699;
XX
DT 02-JUL-2001 (first entry)
XX
DE Human ovarian cancer cell expressed sequence 10799.
XX
KW Human; ovarian cancer; identification; detection; characterisation;
KW tumour; kinase; marker; cytostatic; antisense gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200118542-A2.
XX
PD 15-MAR-2001.
XX
PF 01-SEP-2000; 2000WO-US24199.
XX
PR 03-SEP-1999; 99US-0152547.
PR 16-MAR-2000; 2000US-0190347.
PR 21-MAR-2000; 2000US-0191321.
PR 31-MAY-2000; 2000US-0208382.
PR 20-JUL-2000; 2000US-0220467.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lee J, Thompsho P, Lillie J;
XX
DR WPI; 2001-211428/21.
XX
PT Detection, assessment, prevention and therapy of ovarian cancer,
PT comprises detecting changes in the expression of a variety of markers -
XX
PS Claim 1; Page 1002; 1198pp; English.
XX
CC The present invention describes a method for assessing whether a patient
CC is afflicted with ovarian cancer by comparing: (1) the expression of a
CC marker (I) (see AAF98594 to AAF98730), in a patient sample; and (2) the
CC normal level of expression of (I) in a control non-ovarian cancer
CC sample, where a significant difference between the level of expression
CC in (a) and (b) is an indication that the patient is afflicted with
CC ovarian cancer. (I) have cytostatic activities and can be used in
CC antisense gene therapy. The method, compositions and kits from the
CC present invention can be used for: (1) assessing and treating ovarian
CC cancer; (2) making isolated hybridoma, which produces an antibody useful
CC for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a
CC patient. AAF98573 to AAF98593 represent human kinase marker primers and
CC probes which are used in the exemplification of the present invention.
XX

```
SQ   Sequence 1164 BP; 420 A; 176 C; 195 G; 373 T; 0 other;


  Query Match            89.0%;  Score 97;  DB 22;  Length 1164;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1039 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1098

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db    1099 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1135


RESULT  14
AAX04325
ID   AAX04325 standard; DNA; 1376 BP.
XX
AC   AAX04325;
XX
DT   13-APR-1999  (first entry)
XX
DE   Human secreted protein gene 15 clone HSDES04.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9856804-A1.
XX
PD   17-DEC-1998.
XX
PF   11-JUN-1998;   98WO-US12125.
XX
PR   02-OCT-1997;   97US-0061060.
PR   13-JUN-1997;   97US-0049547.
PR   13-JUN-1997;   97US-0049548.
PR   13-JUN-1997;   97US-0049549.
PR   13-JUN-1997;   97US-0049550.
PR   13-JUN-1997;   97US-0049606.
PR   13-JUN-1997;   97US-0049607.
PR   13-JUN-1997;   97US-0049608.
PR   13-JUN-1997;   97US-0049609.
PR   13-JUN-1997;   97US-0049610.
PR   13-JUN-1997;   97US-0049611.
PR   13-JUN-1997;   97US-0050566.
PR   13-JUN-1997;   97US-0050901.
PR   13-JUN-1997;   97US-0052989.
PR   08-JUL-1997;   97US-0051919.
PR   18-AUG-1997;   97US-0055984.
PR   12-SEP-1997;   97US-0058665.
PR   12-SEP-1997;   97US-0058668.
PR   12-SEP-1997;   97US-0058669.
PR   12-SEP-1997;   97US-0058750.
PR   12-SEP-1997;   97US-0058971.
PR   12-SEP-1997;   97US-0058972.
PR   12-SEP-1997;   97US-0058975.
PR   02-OCT-1997;   97US-0060834.
PR   02-OCT-1997;   97US-0060841.
PR   02-OCT-1997;   97US-0060844.
PR   02-OCT-1997;   97US-0060865.
PR   02-OCT-1997;   97US-0061059.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Brewer LA,  Ebner R,  Ferrie AM,  Feng P,  Greene JM,  Lafleur DW;
PI   Moore PA,  Ni J,  Olsen HS,  Rosen CA,  Ruben SM,  Shi Y,  Young P;
PI   Yu GL;
XX
DR   WPI; 1999-080881/07.
DR   P-PSDB; AAW78140.
XX
PT   New isolated human genes and the secreted polypeptides they encode -
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders
XX
PS   Claim 1; Page 179-180; 380pp; English.
XX
CC   This sequence represents a nucleic acid molecule which encodes a
CC   secreted human protein.  The gene number, and the clone it is derived
CC   from, are detailed in the descriptor line.  The gene can be used to
CC   generate fusion proteins by linking to the gene to a human immunoglobulin
CC   Fc portion (e.g. AAX04302) for increasing the stability of the fused
CC   protein as compared to the human protein only.
CC   The invention relates to 86 novel genes and their fragments (nucleic
CC   acid sequences: AAX04311-X04410; amino acid sequences AAW78126-W78225)
CC   which are useful for preventing, treating or ameliorating  medical
CC   conditions e.g. by protein or gene therapy.  Also, pathological
CC   conditions can be diagnosed by determining the amount of the new
CC   polypeptides in a sample or by determining the presence of mutations in
CC   the new polynucleotides. Specific uses are described for each of the 86
CC   polynucleotides, based on which tissues they are most highly expressed in
CC   (see AAX04311 for described uses).
XX
SQ   Sequence 1376 BP; 499 A; 231 C; 286 G; 359 T; 1 other;


  Query Match            89.0%;  Score 97;  DB 20;  Length 1376;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-16;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1257 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1316

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db    1317 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1353


RESULT  15
AAA27985
ID   AAA27985 standard; cDNA; 1495 BP.
XX
AC   AAA27985;
XX
DT   29-AUG-2000  (first entry)
XX
DE   Corn CCR4 transcription factor nucleotide sequence #2.
XX
KW   CCR4; transcription factor; transgenic plant; antibody production;
KW   corn; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1266
FT                   /*tag= a
FT                   /product= "CCR4"
FT                   /note= "Transcription factor, no start codon given"
XX
PN   WO200032783-A1.
XX
PD   08-JUN-2000.
XX
PF   01-DEC-1999;   99WO-US28334.
XX
```

PR 02-DEC-1998; 98US-0110585.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Allen SM, Sakai H;
XX
DR WPI; 2000-412329/35.
DR P-PSDB; AAY94658.
XX
PT New nucleic acids encoding CCR4 transcription factor in plants and
PT seeds, useful for producing transgenic plants, antibodies and selecting
PT polynucleotides affecting transcription factor polypeptide expression
PT -
XX
PS Claim 2; Page 35-36; 49pp; English.
XX
CC This invention relates to nucleic acid fragments encoding CCR4
CC transcription factors in plants and seeds. The present sequence
CC represents a corn contig encoding CCR4. In saccharomyces cerevisiae the
CC CCR4 gene encodes a transcription factor that is required for the
CC regulation of a number of genes. CCR4 is a component of a multisubunit
CC complex that activates expression of the glucose repressible alcohol
CC dehydrogenase (ADH2) gene. The CCR4 gene suppresses ADH2 expression. The
CC invention includes rice, corn, soybean and wheat CCR4 sequences. The
CC polynucleotide sequences are useful for selecting an isolated
CC polynucleotide that affects the expression levels of transcription factor
CC polypeptide in a plant cell. This can be achieved by introducing the
CC nucleotide sequence into a plant cell and comparing the level of CCR4
CC polypeptide produced, with a plant cell that does not contain the
CC polynucleotide. The polynucleotide sequences are useful as primers and
CC probes for obtaining a nucleic acid fragment encoding a transcription
CC factor polypeptide. An expression cassette comprising the CCR4
CC polynucleotide sequence is useful for positive selection of a transformed
CC cell by transforming it to a plant cell and growing the transformed plant
CC cell and allowing expression of the polynucleotide in an amount
CC sufficient to complement a CCR4 auxotroph to provide a positive selection
CC means. Polypeptides encoded by the nucleotide sequences are useful for
CC producing antibodies. The nucleotide sequences are also useful for
CC producing transgenic plants that express altered levels of CCR4
CC transcription factor, and for mapping the genes encoding CCR4
CC transcription factor.
XX
SQ Sequence 1495 BP; 488 A; 314 C; 318 G; 375 T; 0 other;

Query Match 89.0%; Score 97; DB 21; Length 1495;
Best Local Similarity 100.0%; Pred. No. 1.7e-16;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1378 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1437

Qy 71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||
Db 1438 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1474

Search completed: January 3, 2002, 10:27:59
Job time: 3912 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:59:07 ; Search time 1166.04 Seconds
(without alignments)
1004.502 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 11351937 seqs, 5372889281 residues

Word size : 0

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	102	93.6	301	11	BF726451	BF726451 by06g04.y
	2	102	93.6	405	10	AV682792	AV682792 AV682792
	3	102	93.6	486	10	AV681638	AV681638 AV681638
	4	102	93.6	513	10	AV755973	AV755973 AV755973
	5	102	93.6	694	10	AV734386	AV734386 AV734386
c	6	101	92.7	434	10	AI954183	AI954183 wx80g05.x
	7	101	92.7	562	10	AW636413	AW636413 bl46d05.w
	8	100	91.7	145	10	AU038865	AU038865 AU038865
	9	100	91.7	187	10	AL079963	AL079963 DKFZp586L
	10	100	91.7	383	10	AL045163	AL045163 DKFZp434A
	11	100	91.7	876	12	AF116682	AF116682 Homo sapi
	12	99	90.8	150	10	AU175731	AU175731 AU175731
	13	99	90.8	164	10	AA117464	AA117464 mn22d09.r
	14	99	90.8	180	10	AW598705	AW598705 sj94h09.y
	15	99	90.8	245	10	AA638677	AA638677 vo56b09.r
	16	99	90.8	267	10	AW087445	AW087445 xb98g05.x
c	17	99	90.8	334	10	AI537273	AI537273 tp07a04.x
	18	99	90.8	343	10	BE057182	BE057182 sm97f08.y
	19	99	90.8	354	10	AA575800	AA575800 vn63a02.r
	20	99	90.8	360	10	BE330213	BE330213 so74g08.y
	21	99	90.8	391	10	AV723745	AV723745 AV723745
	22	99	90.8	394	10	AL042745	AL042745 DKFZp434C
	23	99	90.8	476	10	AW096620	AW096620 EST289800
	24	99	90.8	629	11	BF969228	BF969228 602269742
	25	99	90.8	935	11	BE878014	BE878014 601489792
	26	98	89.9	146	10	AA590017	AA590017 vm22e09.r
	27	98	89.9	185	10	BE047852	BE047852 tz43b10.y
	28	98	89.9	229	11	BI142829	BI142829 rk74e11.y
	29	98	89.9	241	10	BE022702	BE022702 sm87d02.y
	30	98	89.9	243	10	AA571665	AA571665 vm09e04.r
c	31	98	89.9	261	10	AW129106	AW129106 xe97d08.x
c	32	98	89.9	276	10	AI249323	AI249323 qx48a07.x
	33	98	89.9	302	10	BE683281	BE683281 182759 MA
	34	98	89.9	318	11	BE800962	BE800962 sr07d02.y
	35	98	89.9	319	10	AW156548	AW156548 se28e08.y
	36	98	89.9	368	10	BE022471	BE022471 sm74d03.y
	37	98	89.9	375	10	AV759831	AV759831 AV759831
	38	98	89.9	403	10	AV681858	AV681858 AV681858
	39	98	89.9	417	10	BE538997	BE538997 601059713
	40	98	89.9	439	11	BF727092	BF727092 by16f02.y
	41	98	89.9	451	10	AV682195	AV682195 AV682195
	42	98	89.9	515	11	BG622615	BG622615 602647359
	43	98	89.9	535	10	BE749464	BE749464 200401 MA
	44	98	89.9	562	11	BG285103	BG285103 602409377
	45	98	89.9	562	11	BG621124	BG621124 602616814

ALIGNMENTS

RESULT 1
BF726451
LOCUS BF726451 301 bp mRNA EST 05-JAN-2001
DEFINITION by06g04.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by06g04 5', mRNA sequence.
ACCESSION BF726451
VERSION BF726451.1 GI:12042362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 301)
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 06 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .301
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="by06g04"
/clone_lib="Human Lens cDNA (Un-normalized, unamplified): BY"
/tissue_type="Lens"
/dev_stage="Adult"
/lab_host="EMDH10B"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies , essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
BASE COUNT 153 a 41 c 33 g 74 t
ORIGIN

Query Match 93.6%; Score 102; DB 11; Length 301;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 10 AAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 69

Qy 66 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||||||||
Db 70 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 111

RESULT 2
AV682792
LOCUS AV682792 405 bp mRNA EST 25-SEP-2000
DEFINITION AV682792 GKB Homo sapiens cDNA clone GKBAGA07 5', mRNA sequence.
ACCESSION AV682792
VERSION AV682792.1 GI:10284655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS Wu,T., Qian,B., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang ,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA GK- clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES Location/Qualifiers
source 1. .405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKBAGA07"
/clone_lib="GKB"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 195 a 57 c 61 g 92 t
ORIGIN

Query Match 93.6%; Score 102; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 1.9e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 33 AAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 92

Qy 66 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||||||||
Db 93 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 134

RESULT 3
AV681638
LOCUS AV681638 486 bp mRNA EST 25-SEP-2000
DEFINITION AV681638 GKB Homo sapiens cDNA clone GKBAJH04 5', mRNA sequence.
ACCESSION AV681638
VERSION AV681638.1 GI:10283501
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Wu,T., Qian,B., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang ,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA GK- clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES Location/Qualifiers
source 1. .486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKBAJH04"
/clone_lib="GKB"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 288 a 78 c 92 g 28 t
ORIGIN

Query Match 93.6%; Score 102; DB 10; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.6e-08;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 aaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 26 AAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 85

Qy 66 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
|||||||||||||||||||||||||||||||||||||||||||
Db 86 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 127

RESULT 4
AV755973
LOCUS AV755973 513 bp mRNA EST 19-OCT-2000
DEFINITION AV755973 BM Homo sapiens cDNA clone BMFAPE03 5', mRNA sequence.
ACCESSION AV755973
VERSION AV755973.1 GI:10913821
KEYWORDS EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
            Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
            ,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
            Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
  TITLE     Homo sapiens cDNA BM clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .513
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BMFAPE03"
                     /clone_lib="BM"
                     /tissue_type="Bone marrow"
                     /cell_type="CD34+ hematopoietic stem/progenitor cell"
                     /lab_host="BM25.8"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      200 a     94 c     91 g    126 t      2 others
ORIGIN


  Query Match             93.6%;  Score 102;  DB 10;  Length 513;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-08;
  Matches  102;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 aaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 AAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 146

Qy     66 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db    147 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188


RESULT   5
AV734386
LOCUS       AV734386      694 bp    mRNA            EST       17-OCT-2000
DEFINITION  AV734386 cdA Homo sapiens cDNA clone cdAAND06 5', mRNA sequence.
ACCESSION   AV734386
VERSION     AV734386.1  GI:10851931
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 694)
  AUTHORS   Yang,Y., Song,H., Peng,Y., Gu,Y., Gao,G., Xiao,H., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
            ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R., Chen,J.,
            Chen,Z. and Han,Z.
  TITLE     Homo sapiens cDNA cdA clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .694
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="cdAAND06"
                     /clone_lib="cdA"
                     /tissue_type="pheochromocytoma"
                     /dev_stage="Adult"
                     /lab_host="BM25.8"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      302 a    127 c     99 g    166 t
ORIGIN


  Query Match             93.6%;  Score 102;  DB 10;  Length 694;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-08;
  Matches  102;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 aaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     19 AAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 78

Qy     66 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db     79 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 120


RESULT   6
AI954183/c
LOCUS       AI954183      434 bp    mRNA            EST       06-SEP-1999
DEFINITION  wx80g05.x1 NCI_CGAP_Ov38 Homo sapiens cDNA clone IMAGE:2550008 3'
            similar to contains element MER22 repetitive element ;, mRNA
            sequence.
ACCESSION   AI954183
VERSION     AI954183.1  GI:5746493
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 434)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeff Struewing, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA
            Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA
            Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 405.
FEATURES             Location/Qualifiers
     source          1. .434
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2550008"
                     /clone_lib="NCI_CGAP_Ov38"
                     /tissue_type="normal epithelium"
                     /lab_host="DH10B"
                     /note="Organ: ovary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies."
BASE COUNT      103 a    105 c     103 g    123 t
ORIGIN


  Query Match             92.7%;  Score 101;  DB 10;  Length 434;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-08;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      7 aaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    113 AAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 54

Qy     67 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db     53 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 13


RESULT   7
AW636413
LOCUS       AW636413      562 bp    mRNA            EST       26-APR-2001
DEFINITION  bl46d05.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
            laevis cDNA clone PBX0046D05 5', mRNA sequence.
ACCESSION   AW636413
VERSION     AW636413.1  GI:7393494
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 562)
  AUTHORS   Blackshear,P.J., Lai,W.S., Thorn,J.M., Kennington,E.A., Staffa,N.G.
            Jr., Moore,D.T., Bouffard,G.G., Beckstrom-Sternberg,S.M., Touchman
            ,J.W., Bonaldo,M.F. and Soares,M.B.
  TITLE     The NIEHS Xenopus maternal EST project: interim analysis of the
            first 13,879 ESTs from unfertilized eggs
  JOURNAL   Gene 267 (1), 71-87 (2001)
  MEDLINE   21211403
COMMENT     Contact: Perry J. Blackshear
            Office of Clinical Research and Laboratory of Signal Transduction
            National Institute of Environmental Health Sciences
            A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
            USA
            Tel: 919 541-4899
            Fax: 919 541-4571
            Email: black009@niehs.nih.gov
            Clone is available through Research Genetics, Inc., 2130 Memorial
            Parkway, Huntsville, AL 35901
             phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
            cdna@resgen.com
             DNA Sequencing and analyses performed by National Institutes of
            Health Intramural Sequencing Center (NISC).
            PCR PRimers
            FORWARD: TGTAAAACGACGGCCAGT
            BACKWARD: CAGGAAACAGCTATGACC
            Plate: 0046  row: D  column: 05
            Seq primer: T7 primer.
FEATURES             Location/Qualifiers
     source          1. .562
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="PBX0046D05"
                     /clone_lib="Blackshear/Soares normalized Xenopus egg
                     library"
                     /sex="female"
                     /tissue_type="unfertilized egg"
                     /cell_type="unfertilized egg"
                     /dev_stage="unfertilized egg"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3-Pac; Site_1: EcoRI; Site_2: NotI;
                     PolyA-selected mRNA was prepared from unfertilized Xenopus
                     laevis eggs.  The library was constructed in the vector
                     pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and
                     Soares, M.B. 'Normalization and subtraction:  two
                     approaches to facilitate gene discovery',  Genome Research
                     6:791-806, 1996.  The first strand synthesis used a
                     NotI-dT18 primer; double stranded cDNAs were ligated to
                     EcoRI adapters, digested with NotI, and directionally
                     cloned into the NotI and EcoRI-digested pT7T3-Pac vector.
                     The library contained approximately 7.2 X 10^5
                     recombinants, with average insert sizes of 1-1.5 kb."
BASE COUNT      225 a     90 c     72 g    175 t
ORIGIN


  Query Match             92.7%;  Score 101;  DB 10;  Length 562;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-08;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 aaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 66
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    411 AAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 470

Qy     67 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db    471 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 511


RESULT   8
AU038865
LOCUS       AU038865      145 bp    mRNA            EST       29-MAR-1999
DEFINITION  AU038865 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSL574, mRNA sequence.
ACCESSION   AU038865
VERSION     AU038865.1  GI:3985618
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 145)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            3-3-10 Ten-nodai, Tsukuba, Ibaraki 305, Japan
            Email: d402hu@sakura.cc.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'.
FEATURES             Location/Qualifiers
     source          1. .145
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSL574"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT      106 a      3 c      0 g     36 t
ORIGIN


  Query Match             91.7%;  Score 100;  DB 10;  Length 145;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-07;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 aaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 66
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     46 AAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 105

Qy     67 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 106
          ||||||||||||||||||||||||||||||||||||||
Db    106 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 145


RESULT   9
AL079963
LOCUS       AL079963      187 bp    mRNA            EST       29-FEB-2000
DEFINITION  DKFZp586L1818_r1 586 (synonym: hutel) Homo sapiens cDNA clone
```

```
            DKFZp586L1818 5', mRNA sequence.
ACCESSION   AL079963
VERSION     AL079963.1  GI:5435539
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 187)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Koehrer, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Koehrer K
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZp586L1818) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .187
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp586L1818"
                     /clone_lib="586 (synonym: hute1)"
                     /tissue_type="uterus"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
BASE COUNT      135 a      8 c     15 g     29 t
ORIGIN


  Query Match             91.7%;  Score 100;  DB 10;  Length 187;
  Best Local Similarity   100.0%;  Pred. No. 9.2e-08;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 aaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 AAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 136

Qy     68 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db    137 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 176


RESULT  10
AL045163
LOCUS       AL045163      383 bp    mRNA            EST       29-FEB-2000
DEFINITION  DKFZp434A1050_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434A1050 5', mRNA sequence.
ACCESSION   AL045163
VERSION     AL045163.2  GI:5935969
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 383)
  AUTHORS   Poustka,A., Klein,M., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Poustka, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     On Jul 9, 1999 this sequence version replaced gi:5433326.
            Contact: Poustka A.J.
            Department Lehrach
            Max-Planck-Institute for Molecular Genetics
            Ihnestrasse 73, 14195 Berlin, Germany
            Tel: +49-30-84131623
            Fax: +49-30-84131128
            Email: poustka@mpimg-berlin-dahlem.mpg.de
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by DKFZ (German Cancer Research Center,
            Heidelberg/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            No s1 sequence available.
            This clone (DKFZp434A1050) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .383
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434A1050"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      164 a      66 c      66 g      85 t       2 others
ORIGIN


  Query Match             91.7%;  Score 100;  DB 10;  Length 383;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-08;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    261 ATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 320

Qy     70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||
Db    321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 360


RESULT  11
AF116682
LOCUS       AF116682      876 bp    mRNA            HTC       08-MAY-2001
DEFINITION  Homo sapiens PRO2013 mRNA, complete cds.
ACCESSION   AF116682
VERSION     AF116682.1  GI:7959862
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 876)
  AUTHORS   Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhou,G., Ouyang,S., Luo,L.,
            Bi,J., Liu,M. and He,F.
  TITLE     Functional prediction of the coding sequences of 121 new genes
            deduced by analysis of cDNA clones from human fetal liver
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 876)
  AUTHORS   Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhou,G., Ouyang,S., Luo,L.,
            Bi,J., Liu,M. and He,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-DEC-1998) Department of Experimental Hematology,
            Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing
            100850, P. R. China
FEATURES             Location/Qualifiers
     source          1. .876
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="FLB7547"
                     /tissue_type="liver"
                     /dev_stage="fetal"
     CDS             136. .381
                     /note="predicted protein of HQ2013"
```

```
                     /codon_start=1
                     /product="PRO2013"
                     /protein_id="AAF71102.1"
                     /db_xref="GI:7959863"
                     /translation="MNVDHEVNLLVEEIHRLGSKNADGKLSVKFGVLFRDDKCANLFE
                     ALVGTLKAAKRRKIVTYPGELLLQGVHDDVDIILLQD"
BASE COUNT      333 a    131 c    167 g    245 t
ORIGIN


  Query Match             91.7%;  Score 100;  DB 12;  Length 876;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-08;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 aaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 67
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    775 AAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 834

Qy     68 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db    835 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 874


RESULT  12
AU175731
LOCUS       AU175731      150 bp    mRNA            EST       05-JUL-2001
DEFINITION  AU175731 Marsupenaeus japonicus adult Marsupenaeus japonicus cDNA
            clone PJA376 reverse, mRNA sequence.
ACCESSION   AU175731
VERSION     AU175731.1  GI:14622138
KEYWORDS    EST.
SOURCE      Marsupenaeus japonicus.
  ORGANISM  Marsupenaeus japonicus
            Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
            Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
            Penaeidae; Marsupenaeus.
REFERENCE   1  (bases 1 to 150)
  AUTHORS   Rojtinnakorn,J., Hirono,I., Aoki,T., Itami,T. and Takahashi,Y.
  TITLE     Gene expression in non-infected and viral infected hemocytes of
            kuruma prawn (Penaeus japonicus)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Takashi Aoki
            Aquatic Biosciences
            Tokyo University of Fisheries
            Konan 4-5-7, Tokyo, Minato-ku 108-8477, Japan
            Tel: 81-3-5463-0689
            Fax: 81-3-5463-0690
            Email: aoki@tokyo-u-fish.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .150
                     /organism="Marsupenaeus japonicus"
                     /db_xref="taxon:27405"
                     /clone="PJA376"
                     /clone_lib="Marsupenaeus japonicus adult"
                     /cell_type="hemocytes"
                     /dev_stage="adult"
                     /note="Vector: lambda ZAP II"
BASE COUNT      118 a      3 c      7 g     21 t      1 others
ORIGIN


  Query Match             90.8%;  Score 99;  DB 10;  Length 150;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-07;
  Matches   99;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 aatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 68
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     43 AATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 102

Qy     69 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||
Db    103 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 141


RESULT  13
AA117464
LOCUS       AA117464      164 bp    mRNA            EST       15-NOV-1996
DEFINITION  mn22d09.r1 Beddington mouse embryonic region Mus musculus cDNA
            clone IMAGE:538673 5', mRNA sequence.
ACCESSION   AA117464
VERSION     AA117464.1  GI:1672477
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 164)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:325609
            Putative full length read
            vector to vector length is 225
            Seq primer: -40m13 ET
            High quality sequence stop: 142.
FEATURES             Location/Qualifiers
     source          1. .164
                     /organism="Mus musculus"
                     /strain="C57BL6 x DBA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:538673"
                     /clone_lib="Beddington mouse embryonic region"
                     /sex="pooled"
                     /tissue_type="embryo"
                     /dev_stage="7.5dpc"
                     /lab_host="DH12S"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT. Gastrulating embryos were collected at 7.5dpc
                     from C57BL6 x DBA matings, excluding embryos that had
                     developed head folds and all extraembryonic tissues.
                     Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
                     Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT      114 a     16 c     16 g     18 t
ORIGIN


  Query Match             90.8%;  Score 99;  DB 10;  Length 164;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-07;
  Matches   99;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     58 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 117

Qy     71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          |||||||||||||||||||||||||||||||||||||||
Db    118 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 156


RESULT  14
AW598705
LOCUS       AW598705      180 bp    mRNA            EST       22-MAR-2000
```

```
DEFINITION  sj94h09.y1 Gm-c1023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1023-2298 5', mRNA sequence.
ACCESSION   AW598705
VERSION     AW598705.1  GI:7286218
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 180)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Putative full length read
            vector to vector length is.
FEATURES             Location/Qualifiers
     source          1. .180
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1023-2298"
                     /clone_lib="Gm-c1023"
                     /tissue_type="seed coats of greenhouse grown plants"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     cDNA library was constructed from mRNA isolated from seed
                     coats (100-200mgs) of greenhouse grown plants.  The
                     library was prepared using the Life Technologies
                     pSuperScript cDNA library construction kit.  Complimentary
                     DNA was synthesized from mRNA using a poly (dT) sequence
                     with a Not I restrictions site.  Sal I linkers adapers
                     were ligated to the blunt-ended cDNA fragments followed by
                     Not I digestion.  The cDNA fragments were directionally
                     cloned into the Not I-Sal I restriction site of the
                     pSPORT1 vector.  The ligated cDNA fragments were
                     transformed into E.coli Electromax DH10B host cells. This
                     library was constructed by Dr. Lila Vodkin and Dr. Anu
                     Khanna."
BASE COUNT      114 a      5 c     20 g     41 t
ORIGIN


  Query Match             90.8%;  Score 99;  DB 10;  Length 180;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-07;
  Matches   99;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 aatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 68
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 AATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 140

Qy     69 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||
Db    141 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 179
```

```
RESULT  15
AA638677
LOCUS       AA638677      245 bp    mRNA            EST       22-OCT-1997
DEFINITION  vo56b09.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA
            clone IMAGE:1053881 5', mRNA sequence.
ACCESSION   AA638677
VERSION     AA638677.1  GI:2562265
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 245)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:585457
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 215.
FEATURES             Location/Qualifiers
     source          1. .245
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1053881"
                     /clone_lib="Barstead mouse irradiated colon MPLRB7"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained
                     from 8 week old mouse.  Colon was harvested 72 hours after
                     irradiation with 1400 Gys.  1st strand cDNA was primed
                     with  a Not I - oligo(dT) primer
                     [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors [AATTCGGATCCTTG], digested with Not I and cloned
                     into the Not I and Eco RI sites of  the modified pT7T3
                     vector. Library constructed by Bob Barstead.  "
BASE COUNT      150 a      28 c      24 g      43 t
ORIGIN


  Query Match             90.8%;  Score 99;  DB 10;  Length 245;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-08;
  Matches   99;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 aatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 68
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 AATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 179

Qy     69 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||
Db    180 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 218


Search completed: January  3, 2002, 10:25:00
Job time: 5153 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:57:37 ; Search time 1342.2 Seconds
(without alignments)
1339.735 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 1472140 seqs, 8248589755 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
1	101	92.7	951	9	HSM801406	AL133560 Homo sapi
2	101	92.7	1843	9	BC009470	BC009470 Homo sapi
3	101	92.7	2506	9	AK026927	AK026927 Homo sapi
4	100	91.7	1827	9	AB070106	AB070106 Macaca fa
5	98	89.9	437	3	AF155369	AF155369 Buthus ma
6	98	89.9	1370	9	AB060826	AB060826 Macaca fa
7	98	89.9	2395	9	AF090903	AF090903 Homo sapi
8	97	89.0	144	6	I89947	I89947 Sequence 26
c 9	97	89.0	173	6	AX185626	AX185626 Sequence
10	97	89.0	311	3	AF159977	AF159977 Buthus ma
c 11	97	89.0	396	6	AX093192	AX093192 Sequence
c 12	97	89.0	597	6	AX187052	AX187052 Sequence
13	97	89.0	617	8	AF323605	AF323605 Manihot e
c 14	97	89.0	621	6	AX187131	AX187131 Sequence
15	97	89.0	641	8	AF266463	AF266463 Manihot e
16	97	89.0	868	8	AF049923	AF049923 Petunia x
17	97	89.0	1117	6	AR123145	AR123145 Sequence
18	97	89.0	1117	8	AF243375	AF243375 Glycine m
19	97	89.0	1164	6	AX098194	AX098194 Sequence
20	97	89.0	1489	9	AB048953	AB048953 Macaca fa
21	97	89.0	1851	9	AK026744	AK026744 Homo sapi
22	97	89.0	1906	9	AK025414	AK025414 Homo sapi
23	97	89.0	1926	9	HSM802868	AL512733 Homo sapi
24	97	89.0	1962	6	AX098509	AX098509 Sequence
25	97	89.0	2013	10	BC006019	BC006019 Mus muscu
26	97	89.0	2032	9	AK027204	AK027204 Homo sapi
27	97	89.0	2463	10	BC005449	BC005449 Mus muscu
28	97	89.0	2844	9	HSM802647	AL353940 Homo sapi
29	97	89.0	2943	10	BC005526	BC005526 Mus muscu
30	96	88.1	132	6	A08913	A08913 H.sapiens (
c 31	96	88.1	310	6	AX187085	AX187085 Sequence
32	96	88.1	353	11	CNS06EJ3	AL395173 T7 end of
33	96	88.1	363	3	AF135821	AF135821 Mesobuthu
34	96	88.1	384	3	AF155365	AF155365 Buthus ma
c 35	96	88.1	396	6	AX093215	AX093215 Sequence
36	96	88.1	440	8	ATP25A	Y11790 A.thaliana
37	96	88.1	750	9	HSM802288	AL137550 Homo sapi
38	96	88.1	801	10	BC005501	BC005501 Mus muscu
39	96	88.1	944	3	BFZ83268	Z83268 B.floridae
40	96	88.1	1022	8	PSPIAA45R	X68215 P.sativum m
41	96	88.1	1252	9	HSSTA	X82434 H.sapiens m
42	96	88.1	1412	3	AF083228	AF083228 Caenorhab
43	96	88.1	1474	6	A65341	A65341 Sequence 64
44	96	88.1	1474	6	AR150479	AR150479 Sequence
45	96	88.1	1555	9	AB063070	AB063070 Macaca fa

ALIGNMENTS

RESULT 1
HSM801406
LOCUS HSM801406 951 bp mRNA PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp434M1414 (from clone DKFZp434M1414); partial cds.
ACCESSION AL133560
VERSION AL133560.1 GI:6599130
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 951)
AUTHORS Koehrer,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
TITLE Direct Submission
JOURNAL Submitted (15-DEC-1999) MIPS, Am Klopferspitz 18a, D-82152 Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical Research Center at the Charite,

Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp434M1414) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.

```
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434M1414"
                     /clone_lib="434 (synonym: htes3). Vector pSport1; host
                     DH10B; sites NotI + SalI"
                     /dev_stage="adult"
                     /tissue_type="testis"
     gene            1. .513
                     /gene="DKFZp434M1414"
     CDS             <1. .513
                     /gene="DKFZp434M1414"
                     /note="KIAA0829, partly"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="CAB63714.1"
                     /db_xref="GI:6599131"
                     /translation="ELIREVEMGPFKHTVDDGLDIRKAAFECMYTLLDSCLDRLDIFE
                     FLNHVEDGLKDHYDIKMLTFLMLVRLSTLCPSAVLQRLDRLVEPLRATCTTKVKANSV
                     KQEFEKQDELKRSAMRAVAALLTIPEAEKSPLMSEFQSQISSNPELAAIFESIQKDSS
                     STNLESMDTS"
     polyA_signal    836. .841
     polyA_site      853
BASE COUNT      365 a    134 c    185 g    267 t
ORIGIN
```

```
  Query Match             92.7%;  Score 101;  DB 9;  Length 951;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-29;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 aaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     847 AAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 906

Qy      67 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||
Db     907 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 947
```

```
RESULT   2
BC009470
LOCUS       BC009470     1843 bp    mRNA            PRI       12-JUL-2001
DEFINITION  Homo sapiens, protein kinase, interferon-inducible double stranded
            RNA dependent activator, clone MGC:1978 IMAGE:3503747, mRNA,
            complete cds.
ACCESSION   BC009470
VERSION     BC009470.1  GI:14495716
KEYWORDS    MGC.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1843)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JUN-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcgsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 6 Row: d Column: 22
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 3290197.
FEATURES             Location/Qualifiers
     source          1. .1843
                     /organism="Homo sapiens"
                     /db_xref="LocusID:8575"
                     /db_xref="taxon:9606"
                     /clone="MGC:1978 IMAGE:3503747"
                     /tissue_type="Brain, neuroblastoma"
                     /clone_lib="NIH_MGC_19"
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     CDS             108. .1049
                     /codon_start=1
                     /product="protein kinase, interferon-inducible double
                     stranded RNA dependent activator"
                     /protein_id="AAH09470.1"
                     /db_xref="GI:14495717"
                     /translation="MSQSRHRAEAPPLEREDSGTFSLGKMITAKPGKTPIQVLHEYGM
                     KTKNIPVYECERSDVQIHVPTFTFRVTVGDITCTGEGTSKKLAKHRAAEAAINILKAN
                     ASICFAVPDPLMPDPSKQPKNQLNPIGSLQELAIHHGWRLPEYTLSQEGGPAHKREYT
                     TICRLESFMETGKGASKKQAKRNAAEKFLAKFSNISPENHISLTNVVGHSLGCTWHSL
                     RNSPGEKINLLKRSLLSIPNTDYIQLLSEIAKEQGFNITYLDIDELSANGQYQCLAEL
                     STSPITVCHGSGISCGNAQSDAAHNALQYLKIIAERK"
BASE COUNT      635 a    366 c    360 g    482 t
ORIGIN
```

```
  Query Match             92.7%;  Score 101;  DB 9;  Length 1843;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-29;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 aaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1717 AAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

Qy      67 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||
Db    1777 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1817
```

```
RESULT   3
AK026927
LOCUS       AK026927     2506 bp    mRNA            PRI       29-SEP-2000
DEFINITION  Homo sapiens cDNA: FLJ23274 fis, clone HEP02623, highly similar to
            AF161354 Homo sapiens HSPC091 mRNA.
ACCESSION   AK026927
VERSION     AK026927.1  GI:10439899
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Homo sapiens hepatoma cell_line:HepG2 cDNA to mRNA, clone_lib:HEP
            clone:HEP02623.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y.,
            Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
```

```
            Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.
  TITLE     NEDO human cDNA sequencing project
  JOURNAL   Unpublished (2000)
REFERENCE   2  (bases 1 to 2506)
  AUTHORS   Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
            Shibahara,T., Tanaka,T. and Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-AUG-2000) to the DDBJ/EMBL/GenBank databases. Sumio
            Sugano, Institute of Medical Science, University of Tokyo,
            Laboratory of Genome Structure Analysis, Human Genome Center;
            Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
            (E-mail:cdna1@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
            Fax:81-3-5449-5416)
COMMENT     NEDO human cDNA sequencing project supported by Ministry of
            International Trade and Industry of Japan; cDNA full insert
            sequencing: Research Association for Biotechnology; cDNA library
            construction, 5'- & 3'-end one pass sequencing: Departent of
            Virology and Human Genome Center, Institute of Medical Science,
            University of Tokyo (partly supported by Science and Technology
            Agency).
FEATURES             Location/Qualifiers
     source          1. .2506
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="HepG2"
                     /cell_type="hepatoma"
                     /clone="HEP02623"
                     /clone_lib="HEP"
                     /note="cloning vector pME18SFL3"
     misc_feature    1. .2506
                     /note="highly similar to AF161354 Homo sapiens HSPC091
                     mRNA"
BASE COUNT      839 a    450 c    474 g    743 t
ORIGIN


  Query Match             92.7%;  Score 101;  DB 9;  Length 2506;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-29;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 aaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 65
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2406 AAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2465

Qy     66 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 106
          ||||||||||||||||||||||||||||||||||||||||||
Db   2466 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2506


RESULT   4
AB070106
LOCUS       AB070106     1827 bp    mRNA            PRI       16-AUG-2001
DEFINITION  Macaca fascicularis testis cDNA clone:QtsA-14022, full insert
            sequence.
ACCESSION   AB070106
VERSION     AB070106.1  GI:15208052
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Macaca fascicularis adult male testis cDNA to mRNA,
            clone_lib:macaque testis cDNA library QtsA clone:QtsA-14022.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 1827)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirai,M.,
            Terao,K. and Sugano,S.
  TITLE     Isolation of novel full-length cDNA clones from macaque testis cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1827)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-AUG-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     Lab host:    TOP10
            Vector:      pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:   DraIII (CACTGTGTG)
            R.  Site2:   DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
            were constructed by oligo-capping method (Sugano et al.,University
            of Tokyo, Institute of Medical
            Science). Custom primer used for sequencing
            ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
            3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1. .1827
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QtsA-14022"
                     /sex="male"
                     /tissue_type="testis"
                     /clone_lib="macaque testis cDNA library QtsA"
                     /dev_stage="adult"
     CDS             132. .1691
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="BAB63051.1"
                     /db_xref="GI:15208053"
                     /translation="MSEMSIDRNIHHNLSPGIPISVQTEESWLQDLLGKVQSRKKTAS
                     KVKKAEPQCLASKLREKWVINPEEPKLNILYELEFKEDFITLFEPSLRTLPSIGPPTI
                     LAYKEESSNLDINFKDEEEETSPKCEFCGSDLRAFFSNVDVCSESKGHASCCIAFQNL
                     IDYIYEEQIKIKSPKAELIAIDPHAAHYSEVDRLKAKEKALQRKQEQRMARHFAIISR
                     EQTHFSEDDSKRLKTISYQLSVDIPEKQTIDDIVFDFQLRNSNMSIICCDSRIACGKV
                     VRNELLEKHYKHGSKFLTSFPDGTTQIFYPSGNLAIIRVPNKVNGFTCIVQEDMPTNP
                     AILAVLDSSGRSSCYHPNGNVWVYINILGGQYSDQAGNRIRAWNWSNSITSSPFASFK
                     PVFLALNRYIGVRILEQDKISITFLAMGQQARISVGTKVKLPNPEEIPVLRYVSGDDL
                     LLLASLIKIRRLFHKLEGCVNFPSSQVWEKLKQPSYLSSLTLKLIALCRSSGIKQDTM
                     KTIRNIINEEI"
BASE COUNT      702 a    350 c    332 g    443 t
ORIGIN


  Query Match             91.7%;  Score 100;  DB 9;  Length 1827;
  Best Local Similarity   100.0%;  Pred. No. 3.4e-29;
  Matches  100;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 aaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 67
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1684 AAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1743

Qy     68 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          ||||||||||||||||||||||||||||||||||||||||||
Db   1744 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1783


RESULT   5
AF155369
LOCUS       AF155369     437 bp    mRNA            INV       13-JUN-2001
DEFINITION  Buthus martensii putative potassium ion channel blocker TXKs1 mRNA,
            complete cds.
ACCESSION   AF155369
VERSION     AF155369.1  GI:14388602
KEYWORDS    .
SOURCE      Chinese scorpion.
  ORGANISM  Buthus martensii
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
            Buthoidea; Buthidae; Buthus.
```

```
REFERENCE   1  (bases 1 to 437)
  AUTHORS   Zhu,S. and Li,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-JUN-1999) Virology Department, Wuhan University,
            Luojia Mountain, Wuhan, Hubei Province 430072, China
FEATURES             Location/Qualifiers
     source          1. .437
                     /organism="Buthus martensii"
                     /db_xref="taxon:34649"
                     /tissue_type="venom gland"
     CDS             56. .238
                     /codon_start=1
                     /product="putative potassium ion channel blocker TXKs1"
                     /protein_id="AAK61819.1"
                     /db_xref="GI:14388603"
                     /translation="MNRLTTIILMLIVINVIMDDISESKVAAGIVCKVCKIICGMQGK
                     KVNICKAPIKCKCKKG"
BASE COUNT      261 a     36 c     49 g     91 t
ORIGIN


  Query Match             89.9%;  Score 98;  DB 3;  Length 437;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-28;
  Matches   98;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     278 ATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 337

Qy      70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           ||||||||||||||||||||||||||||||||||||
Db     338 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 375


RESULT   6
AB0608..6
LOCUS       AB060826     1370 bp    mRNA            PRI       13-JUN-2001
DEFINITION  Macaca fascicularis brain cDNA clone:QtrA-10210, full insert
            sequence.
ACCESSION   AB060826
VERSION     AB060826.1  GI:13874461
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Macaca fascicularis adult male temporal lobe right cDNA to mRNA,
            clone_lib:macaque brain cDNA library QtrA clone:QtrA-10210.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (sites)
  AUTHORS   Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,
            Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.
  TITLE     Isolation of full-length cDNA clones from macaque brain cDNA
            libraries
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1370)
  AUTHORS   Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     Lab host:    TOP10
            Vector:      pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:   DraIII (CACTGTGTG)
            R.  Site2:   DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The Sfi1-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
            were constructed by oligo-capping method
            (Sugano et al., , Institute of Medical Science, University of
            Tokyo).
            Custom primer used for sequencing
            ( 5' end primer [CTTCTGCTCTAAAAGCTGCG];
            3' end primer [CGACCTGCAGCTCGAGCACA] ).
FEATURES             Location/Qualifiers
     source          1. .1370
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
                     /clone="QtrA-10210"
                     /sex="male"
                     /tissue_type="temporal lobe right"
                     /clone_lib="macaque brain cDNA library QtrA"
                     /dev_stage="adult"
BASE COUNT      410 a     281 c     336 g     343 t
ORIGIN


  Query Match             89.9%;  Score 98;  DB 9;  Length 1370;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-28;
  Matches   98;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1266 ATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1325

Qy      70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           ||||||||||||||||||||||||||||||||||||
Db    1326 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1363


RESULT   7
AF090903
LOCUS       AF090903     2395 bp    mRNA            PRI       12-JAN-2000
DEFINITION  Homo sapiens clone HQ0214 PRO0214 mRNA, complete cds.
ACCESSION   AF090903
VERSION     AF090903.1  GI:6690182
KEYWORDS    FLI_CDNA.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2395)
  AUTHORS   Yu,Y., Zhang,C., Luo,L., Ouyang,S., Zhang,S., Li,W., Wu,J.,
            Zhou,S., Liu,M. and He,F.
  TITLE     Functional prediction of the coding sequences of 50 new genes
            deduced by analysis of cDNA clones from human fetal liver
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2395)
  AUTHORS   Yu,Y., Zhang,C., Luo,L., Ouyang,S., Zhang,S., Li,W., Wu,J.,
            Zhou,S., Liu,M. and He,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-1998) Dept. of Experimental Hematology, Beijing
            Institute of Radiation Medcine, 27 Taiping RD, Beijing 100850,
            P.R.China
FEATURES             Location/Qualifiers
     source          1. .2395
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HQ0214"
                     /tissue_type="liver"
                     /dev_stage="fetus"
     CDS             1401. .1649
                     /codon_start=1
                     /product="PRO0214"
                     /protein_id="AAF24027.1"
                     /db_xref="GI:6690183"
                     /translation="MYLSVFKAIKIHWAFCHHVKMDTAWYALITFPFLILTVIAVFAL
                     AVLFYTQKSGKVLNSILPPKQNKIHGLGYNSEVFSTYF"
BASE COUNT      878 a     435 c     364 g     718 t
ORIGIN
```

```
  Query Match             89.9%;  Score 98;  DB 9;  Length 2395;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-28;
  Matches   98;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 atttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 69
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2261 ATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2320

Qy      70 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa.107
           ||||||||||||||||||||||||||||||||||||||
Db    2321 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2358


RESULT   8
I89947
LOCUS       I89947        144 bp    DNA              PAT       10-AUG-1998
DEFINITION  Sequence 26 from patent US 5723315.
ACCESSION   I89947
VERSION     I89947.1  GI:3409887
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 144)
  AUTHORS   Jacobs,K., McCoy,J.M., LaVallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M. and Spaulding,V.
  TITLE     Secreted proteins and polynucleotides encoding them
  JOURNAL   Patent: US 5723315-A 26 03-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .144
                     /organism="unknown"
BASE COUNT      132 a      0 c      0 g      8 t      4 others
ORIGIN


  Query Match             89.0%;  Score 97;  DB 6;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-28;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      20 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 79

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db      80 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 116


RESULT   9
AX185626/c
LOCUS       AX185626      173 bp    DNA              PAT       06-AUG-2001
DEFINITION  Sequence 1321 from Patent WO0142467.
ACCESSION   AX185626
VERSION     AX185626.1  GI:15137029
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 173)
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 1321 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .173
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       14 a      11 c      10 g     137 t       1 others
ORIGIN


  Query Match             89.0%;  Score 97;  DB 6;  Length 173;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-28;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     163 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 104

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db     103 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 67


RESULT  10
AF159977
LOCUS       AF159977      311 bp    mRNA             INV       13-JUN-2001
DEFINITION  Buthus martensii toxin TLP3 mRNA, partial cds.
ACCESSION   AF159977
VERSION     AF159977.1  GI:14388612
KEYWORDS    .
SOURCE      Chinese scorpion.
  ORGANISM  Buthus martensii
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
            Buthoidea; Buthidae; Buthus.
REFERENCE   1  (bases 1 to 311)
  AUTHORS   Shunyi,Z. and Wenxin,L.
  TITLE     cDNA encoding a toxin-like peptide from BmK
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 311)
  AUTHORS   Wenxin,L. and Shunyi,Z.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-JUN-1999) Virology Department, Wuhan University,
            Luojia Mountain, Wuhan, Hubei 430072, China
FEATURES             Location/Qualifiers
     source          1. .311
                     /organism="Buthus martensii"
                     /db_xref="taxon:34649"
                     /tissue_type="venom gland"
     CDS             <1. .142
                     /codon_start=2
                     /product="toxin TLP3"
                     /protein_id="AAK61824.1"
                     /db_xref="GI:14388613"
                     /translation="ENLGEDCENLCKQQKATDGFCRQPHCFCTDMPDNYATRPDTVDP
                     IM"
BASE COUNT      181 a      29 c      33 g      68 t
ORIGIN


  Query Match             89.0%;  Score 97;  DB 3;  Length 311;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-28;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     208 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 267

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db     268 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 304


RESULT  11
AX093192/c
LOCUS       AX093192      396 bp    DNA              PAT       30-MAR-2001
DEFINITION  Sequence 10 from Patent WO0118046.
ACCESSION   AX093192
VERSION     AX093192.1  GI:13509641
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Xu,J. and Stolk,J.A.
  TITLE     Ovarian tumor sequences and methods of use therefor
  JOURNAL   Patent: WO 0118046-A 10 15-MAR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1. .396
                     /note="n = A,T,C or G"
BASE COUNT       85 a     18 c     55 g    153 t     85 others
ORIGIN


  Query Match             89.0%;  Score 97;  DB 6;  Length 396;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-28;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     107 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 48

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||
Db      47 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11


RESULT  12
AX187052/c
LOCUS       AX187052      597 bp    DNA             PAT       06-AUG-2001
DEFINITION  Sequence 2747 from Patent WO0142467.
ACCESSION   AX187052
VERSION     AX187052.1  GI:15138496
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 597)
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 2747 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .597
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      259 a     54 c     39 g    244 t      1 others
ORIGIN


  Query Match             89.0%;  Score 97;  DB 6;  Length 597;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-28;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     141 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 82

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||
Db      81 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 45


RESULT  13
AF323605
LOCUS       AF323605      617 bp    mRNA            PLN       22-FEB-2001
DEFINITION  Manihot esculenta RNA polymerase I, II and III subunit RPB8 mRNA,
            partial cds.
ACCESSION   AF323605
VERSION     AF323605.1  GI:13094964
KEYWORDS    .
SOURCE      cassava.
  ORGANISM  Manihot esculenta
            Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
            Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
            Malpighiales; Euphorbiaceae; Manihot.
REFERENCE   1  (bases 1 to 617)
  AUTHORS   Reilly,K., Winter,M., Han,Y. and Beeching,J.
  TITLE     Cassava RNA polymerase I, II and III subunit with homology to yeast
            RPB8
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 617)
  AUTHORS   Reilly,K., Winter,M., Han,Y. and Beeching,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-NOV-2000) Biology and Biochemistry, University of
            Bath, Claverton Down, Bath BA2 7AY, UK
FEATURES             Location/Qualifiers
     source          1. .617
                     /organism="Manihot esculenta"
                     /db_xref="taxon:3983"
     CDS             <1. .335
                     /note="polI, polII, polIII common subunit"
                     /codon_start=3
                     /product="RNA polymerase I, II and III subunit RPB8"
                     /protein_id="AAK12101.1"
                     /db_xref="GI:13094965"
                     /translation="DMFMHLDVNTEIYPMAVGDKFTMALAHTLNLDGTPDTGYYIQGA
                     RKTLADKYEYIMHGKLYKISEEGSGKAVKAEIYVSYGGLLMMLKGDPSHVSHFELDQR
                     LFLLIRKL"
BASE COUNT      252 a     87 c    110 g    168 t
ORIGIN


  Query Match             89.0%;  Score 97;  DB 8;  Length 617;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-28;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     519 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 578

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||
Db     579 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 615


RESULT  14
AX187131/c
LOCUS       AX187131      621 bp    DNA             PAT       06-AUG-2001
DEFINITION  Sequence 2826 from Patent WO0142467.
ACCESSION   AX187131
VERSION     AX187131.1  GI:15138579
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 2826 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      273 a     63 c     47 g    238 t
ORIGIN


  Query Match             89.0%;  Score 97;  DB 6;  Length 621;
```

Best Local Similarity 100.0%; Pred. No. 4.5e-28;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     135 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 76

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db      75 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 39
```

RESULT 15
AF266463
LOCUS AF266463 641 bp mRNA PLN 24-MAY-2001
DEFINITION Manihot esculenta RNA polymerase subunit mRNA, complete cds.
ACCESSION AF266463
VERSION AF266463.1 GI:14193246
KEYWORDS .
SOURCE cassava.
ORGANISM Manihot esculenta
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Malpighiales; Euphorbiaceae; Manihot.
REFERENCE 1 (bases 1 to 641)
AUTHORS Reilly,K., Winter,M.J., Han,Y., Tohme,J. and Beeching,J.R.
TITLE RNA polymerase subunit isolated from cassava storage roots
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 641)
AUTHORS Reilly,K., Winter,M.J., Han,Y., Tohme,J. and Beeching,J.R.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2000) Biology and Biochemistry, University of
Bath, Claverton Down, Bath BA2 7AY, UK
FEATURES Location/Qualifiers
source 1. .641
/organism="Manihot esculenta"
/db_xref="taxon:3983"
CDS 69. .359
/codon_start=1
/product="RNA polymerase subunit"
/protein_id="AAK55847.1"
/db_xref="GI:14193247"
/translation="MAVGDKFTMALAHTLNLDGTPDTGYYIQGARKTLADKYEYIMHG
KLYKISEEGSGKAVKAEIYVSYGGLLMMLKGDPSHVSHFELDQRLFLLIRKL"
BASE COUNT 257 a 92 c 116 g 176 t
ORIGIN

Query Match 89.0%; Score 97; DB 8; Length 641;
Best Local Similarity 100.0%; Pred. No. 4.5e-28;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 602

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||
Db     603 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 639
```

Search completed: January 3, 2002, 10:05:11
Job time: 4054 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 38.18 Seconds
(without alignments)
646.572 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	102.6	94.1	1051	4	US-09-245-041-10	Sequence 10, Appl
	2	102.6	94.1	1733	3	US-09-073-569-1	Sequence 1, Appli
c	3	102.6	94.1	9589	1	US-07-925-695-1	Sequence 1, Appli
c	4	102.6	94.1	9589	1	US-07-925-695-2	Sequence 2, Appli
	5	101	92.7	111	4	US-09-297-535-23	Sequence 23, Appl
	6	101	92.7	121	4	US-09-297-535-20	Sequence 20, Appl
	7	100.6	92.3	117	1	US-08-702-344-3	Sequence 3, Appli
	8	100.6	92.3	144	1	US-08-702-344-26	Sequence 26, Appl
	9	100.6	92.3	240	1	US-08-628-417-6	Sequence 6, Appli
	10	100.6	92.3	1813	4	US-09-071-224-3	Sequence 3, Appli
	11	100.6	92.3	2082	2	US-08-785-310A-2	Sequence 2, Appli
	12	100.6	92.3	2447	2	US-09-014-969-14	Sequence 14, Appl
	13	100.6	92.3	6671	1	US-08-280-443-1	Sequence 1, Appli
	14	100.6	92.3	6671	1	US-08-457-459-1	Sequence 1, Appli
	15	100.6	92.3	6671	1	US-08-555-678-1	Sequence 1, Appli
	16	100.6	92.3	6671	5	PCT-US95-02275-1	Sequence 1, Appli
	17	100.2	91.9	2184	4	US-08-955-918C-1	Sequence 1, Appli
	18	98.6	90.5	1117	4	US-09-247-373B-33	Sequence 33, Appl
	19	98.2	90.1	1066	1	US-08-157-101A-4	Sequence 4, Appli
	20	97.6	89.5	140	1	US-08-628-417-5	Sequence 5, Appli
	21	96.6	88.6	1882	4	US-09-370-253-1	Sequence 1, Appli
	22	96	88.1	1474	4	US-08-821-994-64	Sequence 64, Appl
	23	95.8	87.9	1641	1	US-08-300-903A-8	Sequence 8, Appli
	24	95	87.2	1582	3	US-08-545-196B-10	Sequence 10, Appl
	25	95	87.2	1582	3	US-08-545-196B-12	Sequence 12, Appl
	26	94	86.2	98	1	US-08-088-658-42	Sequence 42, Appl
	27	94	86.2	98	2	US-08-471-907A-42	Sequence 42, Appl
	28	93.2	85.5	2671	6	5168051-9	Patent No. 5168051
	29	92.8	85.1	790	4	US-09-363-970-4	Sequence 4, Appli
	30	91.8	84.2	1098	3	US-09-248-335-35	Sequence 35, Appl
c	31	91.6	84.0	356	2	US-08-520-678A-22	Sequence 22, Appl
c	32	91.6	84.0	356	4	US-08-897-126-22	Sequence 22, Appl
c	33	91.6	84.0	9646	3	US-08-811-566-1	Sequence 1, Appli
c	34	91.6	84.0	12980	3	US-08-811-566-5	Sequence 5, Appli
	35	91	83.5	1411	4	US-08-964-127-5	Sequence 5, Appli
	36	90.8	83.3	903	5	PCT-US95-06406A-21	Sequence 21, Appl
	37	89.8	82.4	1134	3	US-09-248-335-29	Sequence 29, Appl
c	38	89.4	82.0	260	2	US-08-520-678A-29	Sequence 29, Appl
c	39	89.4	82.0	260	4	US-08-897-126-29	Sequence 29, Appl
	40	89	81.7	1817	1	US-08-473-981A-5	Sequence 5, Appli
	41	89	81.7	1817	2	US-08-474-087-5	Sequence 5, Appli
	42	88.6	81.3	635	1	US-08-455-633A-35	Sequence 35, Appl
	43	88.6	81.3	635	1	US-08-416-336-5	Sequence 5, Appli
	44	88.6	81.3	635	2	US-08-456-460C-35	Sequence 35, Appl
	45	88.6	81.3	635	5	PCT-US94-05354-35	Sequence 35, Appl

ALIGNMENTS

```
RESULT   1
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
;  APPLICANT: Moore, K.
;  APPLICANT: Nagle, D.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
;  TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
;  FILE REFERENCE: 7853-136
;  CURRENT APPLICATION NUMBER: US/09/245,041
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/093,630
;  EARLIER FILING DATE: 1998-07-21
;  EARLIER APPLICATION NUMBER: 60/104,978
;  EARLIER FILING DATE: 1998-10-20
;  NUMBER OF SEQ ID NOS: 131
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 1051
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-245-041-10

  Query Match             94.1%;  Score 102.6;  DB 4;  Length 1051;
  Best Local Similarity   96.3%;  Pred. No. 1.7e-07;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 91

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 140


RESULT   2
US-09-073-569-1
; Sequence 1, Application US/09073569
; Patent No. 6084088
;  GENERAL INFORMATION:
;    APPLICANT:  Sheppard, Paul O.
;    APPLICANT:  Grossmann, Angelika
;    TITLE OF INVENTION:  NOVEL TUMOR ANTIGENS
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZymoGenetics, Inc.
;      STREET:  1201 Eastlake Avenue East
```

```
;      CITY:  Seattle
;      STATE:  WA
;      COUNTRY:  USA
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/073,569
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sawislak, Deborah A
;      REGISTRATION NUMBER:  37,438
;      REFERENCE/DOCKET NUMBER:  97-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6672
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1733 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  34...1344
;      OTHER INFORMATION:
US-09-073-569-1


  Query Match             94.1%;  Score 102.6;  DB 3;  Length 1733;
  Best Local Similarity   96.3%;  Pred. No. 1.5e-07;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1617 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1676

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1677 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 1725


RESULT   3
US-07-925-695-1/c
; Sequence 1, Application US/07925695
; Patent No. 5428145
;  GENERAL INFORMATION:
;    APPLICANT:  OKAMOTO, Hiroaki
;    APPLICANT:  NAKAMURA, Tetsuo
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOME,
;    TITLE OF INVENTION:  POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
;    TITLE OF INVENTION:  DETECTION SYSTEMS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Beveridge, DeGrandi, Weilacher & Young
;      STREET:  1850 M Street, N.W., Suite 800
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/925,695
;      FILING DATE:  19920807
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 287402/91
;      FILING DATE:  09-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 360441/91
;      FILING DATE:  05-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weilacher, Robert G.
;      REGISTRATION NUMBER:  20,531
;      REFERENCE/DOCKET NUMBER:  06/87-48009
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2811
;      TELEFAX:  (202) 659-1462
;      TELEX:  WUI 64470
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9589 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-925-695-1


  Query Match             94.1%;  Score 102.6;  DB 1;  Length 9589;
  Best Local Similarity   96.3%;  Pred. No. 8.9e-08;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480


RESULT   4
US-07-925-695-2/c
; Sequence 2, Application US/07925695
; Patent No. 5428145
;  GENERAL INFORMATION:
;    APPLICANT:  OKAMOTO, Hiroaki
;    APPLICANT:  NAKAMURA, Tetsuo
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOME,
;    TITLE OF INVENTION:  POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
;    TITLE OF INVENTION:  DETECTION SYSTEMS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Beveridge, DeGrandi, Weilacher & Young
;      STREET:  1850 M Street, N.W., Suite 800
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/925,695
;      FILING DATE:  19920807
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 287402/91
;      FILING DATE:  09-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 360441/91
;      FILING DATE:  05-DEC-1991
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weilacher, Robert G.
;      REGISTRATION NUMBER:  20,531
;      REFERENCE/DOCKET NUMBER:  06/87-48009
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2811
;      TELEFAX:  (202) 659-1462
;      TELEX:  WUI 64470
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9589 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-925-695-2


  Query Match             94.1%;  Score 102.6;  DB 1;  Length 9589;
  Best Local Similarity   96.3%;  Pred. No. 8.9e-08;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480


RESULT   5
US-09-297-535-23
; Sequence 23, Application US/09297535
; Patent No. 6268199
; GENERAL INFORMATION:
;  APPLICANT: Meulenberg, Johanna J.M.
;  APPLICANT: Pol, Johannes M.A.
;  APPLICANT: Bos-de Ruijter, Judy N.A.
;  TITLE OF INVENTION: Infectious clones of RNA viruses and vaccines and
;  TITLE OF INVENTION: diagnostic assays derived thereof.
;  FILE REFERENCE: P20749US00
;  CURRENT APPLICATION NUMBER: US/09/297,535
;  CURRENT FILING DATE: 1999-10-12
;  EARLIER APPLICATION NUMBER: PCT/NL97/00593
;  EARLIER FILING DATE: 1997-10-29
;  EARLIER APPLICATION NUMBER: 96203024.3
;  EARLIER FILING DATE: 1996-10-30
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 111
;   TYPE: RNA
;   ORGANISM: Lelystad virus
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: (1)..(111)
;   OTHER INFORMATION: /note="3' end"
US-09-297-535-23


  Query Match             92.7%;  Score 101;  DB 4;  Length 111;
  Best Local Similarity   95.4%;  Pred. No. 5.2e-07;
  Matches  104;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 62

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     63 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacg 111


RESULT   6
US-09-297-535-20
; Sequence 20, Application US/09297535
; Patent No. 6268199
; GENERAL INFORMATION:
;  APPLICANT: Meulenberg, Johanna J.M.
;  APPLICANT: Pol, Johannes M.A.
;  APPLICANT: Bos-de Ruijter, Judy N.A.
;  TITLE OF INVENTION: Infectious clones of RNA viruses and vaccines and
;  TITLE OF INVENTION: diagnostic assays derived thereof.
;  FILE REFERENCE: P20749US00
;  CURRENT APPLICATION NUMBER: US/09/297,535
;  CURRENT FILING DATE: 1999-10-12
;  EARLIER APPLICATION NUMBER: PCT/NL97/00593
;  EARLIER FILING DATE: 1997-10-29
;  EARLIER APPLICATION NUMBER: 96203024.3
;  EARLIER FILING DATE: 1996-10-30
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;   LENGTH: 121
;   TYPE: DNA
;   ORGANISM: Lelystad virus
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: (1)..(121)
;   OTHER INFORMATION: /note="3'end"
US-09-297-535-20


  Query Match             92.7%;  Score 101;  DB 4;  Length 121;
  Best Local Similarity   95.4%;  Pred. No. 5.1e-07;
  Matches  104;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 62

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db     63 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacg 111


RESULT   7
US-08-702-344-3
; Sequence 3, Application US/08702344
; Patent No. 5723315
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/702,344
;      FILING DATE:
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 1164.61 Seconds
(without alignments)
1005.736 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estfun:*
2: em_esthum:*
3: em_estin:*
4: em_estom:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_htc:*
10: gb_est1:*
11: gb_est2:*
12: gb_htc:*
13: gb_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	105.4	96.7	301	11	BF726451	BF726451 by06g04.y
	2	105.4	96.7	405	10	AV682792	AV682792 AV682792
	3	105.4	96.7	486	10	AV681638	AV681638 AV681638
	4	105.4	96.7	513	10	AV755973	AV755973 AV755973
	5	105	96.3	576	13	CNS0256N	AL181688 Tetraodon
c	6	104.6	96.0	340	10	AL513755	AL513755 AL513755
	7	104.6	96.0	363	10	AL514473	AL514473 AL514473
	8	104.2	95.6	200	10	AW185559	AW185559 se81e02.y
c	9	104.2	95.6	449	10	AL514879	AL514879 AL514879
	10	104.2	95.6	1042	11	BG115502	BG115502 602316663
c	11	104.2	95.6	1435	11	BF984813	BF984813 602308063
	12	103.8	95.2	134	13	CNS03F8P	AL241378 Tetraodon
	13	103.8	95.2	142	10	AA117240	AA117240 mn19f12.r
	14	103.8	95.2	147	10	AI340603	AI340603 tb31g08.x
	15	103.8	95.2	267	10	AW087445	AW087445 xb98g05.x
	16	103.8	95.2	386	10	AV759054	AV759054 AV759054
	17	103.8	95.2	402	10	AV682849	AV682849 AV682849
	18	103.8	95.2	415	10	BE689881	BE689881 uw61c06.y
c	19	103.8	95.2	449	10	AL513999	AL513999 AL513999
	20	103.8	95.2	504	10	AV758806	AV758806 AV758806
	21	103.8	95.2	542	10	AV757028	AV757028 AV757028
	22	103.8	95.2	578	10	AW706506	AW706506 sj57c06.y
	23	103.8	95.2	683	10	AV682476	AV682476 AV682476
	24	103.8	95.2	694	10	AV734386	AV734386 AV734386
	25	103.8	95.2	712	10	AV734318	AV734318 AV734318
	26	103.8	95.2	735	10	AV681769	AV681769 AV681769
	27	103.8	95.2	762	10	AV758110	AV758110 AV758110
	28	103.8	95.2	773	10	AA528822	AA528822 nf52d11.s
c	29	103.8	95.2	803	11	BG335466	BG335466 602403936
c	30	103.8	95.2	838	10	AL514075	AL514075 AL514075
	31	103.8	95.2	935	11	BE878014	BE878014 601489792
c	32	103.8	95.2	1016	13	CNS003S9	AL065002 Drosophil
c	33	103.8	95.2	1166	11	BE964682	BE964682 601658218
c	34	103.4	94.9	400	10	AL513741	AL513741 AL513741
c	35	103.4	94.9	626	10	AL513937	AL513937 AL513937
c	36	103.2	94.7	147	10	AV610565	AV610565 AV610565
c	37	103.2	94.7	204	11	BI024203	BI024203 CM3-MT035
	38	103.2	94.7	273	10	AW164794	AW164794 se77h06.y
c	39	103	94.5	956	13	CNS014V0	AL104646 Drosophil
c	40	102.8	94.3	120	10	AI284517	AI284517 qt88g07.x
	41	102.8	94.3	264	10	AA607075	AA607075 vm95h03.r
	42	102.8	94.3	278	10	AA638755	AA638755 vm94h11.r
c	43	102.8	94.3	329	10	AL514627	AL514627 AL514627
	44	102.8	94.3	383	10	AL045163	AL045163 DKFZp434A
c	45	102.8	94.3	408	10	AL514793	AL514793 AL514793

ALIGNMENTS

RESULT 1
BF726451
LOCUS BF726451 301 bp mRNA EST 05-JAN-2001
DEFINITION by06g04.y1 Human Lens cDNA (Un-normalized, unamplified): BY Homo sapiens cDNA clone by06g04 5', mRNA sequence.
ACCESSION BF726451
VERSION BF726451.1 GI:12042362
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 301)
AUTHORS Wistow,G.J., Bernstein,S., Behal,A. and Smith,D.
TITLE NEIBANK: EST analysis and bioinformatics for ocular genomics
JOURNAL Invest. Ophthalmol. Vis. Sci. 41 (2000) In press
COMMENT Contact: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
Plate: 06 row: g column: 04
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .301
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="by06g04"
/clone_lib="Human Lens cDNA (Un-normalized, unamplified): BY"
/tissue_type="Lens"
/dev_stage="Adult"
/lab_host="EMDH10B"

/note="Organ: Eye; Vector: pCMVSPORT6; Two human lenses from different adults (both approximately 40 years old) together yielded 20ug of total RNA and 150ng mRNA for cDNA library synthesis. A directionally cloned cDNA library in the pCMVSPORT6 vector was constructed at Life Technologies , essentially following the protocols of the SuperScript Plasmid System full details of which are contained in the manufacturer's Instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adapter [5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCC(T)15-3']. Not I/blunt end inserts were cloned into the Not I/EcoR V sites in the vector. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."

BASE COUNT 153 a 41 c 33 g 74 t
ORIGIN

Query Match 96.7%; Score 105.4; DB 11; Length 301;
Best Local Similarity 99.1%; Pred. No. 3.8;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      5 AAAAAAAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 64

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||
Db     65 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 111
```

RESULT 2
AV682792
LOCUS AV682792 405 bp mRNA EST 25-SEP-2000
DEFINITION AV682792 GKB Homo sapiens cDNA clone GKBAGA07 5', mRNA sequence.
ACCESSION AV682792
VERSION AV682792.1 GI:10284655
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 405)
AUTHORS Wu,T., Qian,B., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang ,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA GK- clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES Location/Qualifiers
source 1. .405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKBAGA07"
/clone_lib="GKB"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 195 a 57 c 61 g 92 t
ORIGIN

Query Match 96.7%; Score 105.4; DB 10; Length 405;
Best Local Similarity 99.1%; Pred. No. 3.1;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     28 AAAAAAAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 87

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||
Db     88 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 134
```

RESULT 3
AV681638
LOCUS AV681638 486 bp mRNA EST 25-SEP-2000
DEFINITION AV681638 GKB Homo sapiens cDNA clone GKBAJH04 5', mRNA sequence.
ACCESSION AV681638
VERSION AV681638.1 GI:10283501
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Wu,T., Qian,B., Huang,Q., Huang,C., Kang,B., Gao,X., Xu,Z., Xiao,H., Xu,X., Li,N., Peng,Y., Liu,F., Qu,J., Song,H., Cheng,Z., Zeng,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Yang ,Y., Gu,Y., Chen,Z. and Han,Z.
TITLE Homo sapiens cDNA GK- clones
JOURNAL Unpublished (2000)
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES Location/Qualifiers
source 1. .486
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GKBAJH04"
/clone_lib="GKB"
/tissue_type="hepatocellular carcinoma"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 288 a 78 c 92 g 28 t
ORIGIN

Query Match 96.7%; Score 105.4; DB 10; Length 486;
Best Local Similarity 99.1%; Pred. No. 2.7;
Matches 106; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AAAAAAAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 80

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||
Db     81 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 127
```

RESULT 4
AV755973
LOCUS AV755973 513 bp mRNA EST 19-OCT-2000
DEFINITION AV755973 BM Homo sapiens cDNA clone BMFAPE03 5', mRNA sequence.
ACCESSION AV755973
VERSION AV755973.1 GI:10913821
KEYWORDS EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Gu,J., Zhao,M., Huang,Q., Xu,X., Li,Y., Peng,Y., Song,H., Xiao,H.,
            Gu,Y., Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng
            ,L., Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G.,
            Yang,Y., Gao,G., Wang,Z., Zhang,Q., Chen,S., Han,Z. and Chen,Z.
  TITLE     Homo sapiens cDNA BM clones
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@chgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES             Location/Qualifiers
     source          1. .513
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="BMFAPE03"
                     /clone_lib="BM"
                     /tissue_type="Bone marrow"
                     /cell_type="CD34+ hematopoietic stem/progenitor cell"
                     /lab_host="BM25.8"
                     /note="Vector: pTriplEx2; Site_1: sfiIA; Site_2: sfiIB"
BASE COUNT      200 a      94 c      91 g     126 t       2 others
ORIGIN


  Query Match             96.7%;  Score 105.4;  DB 10;  Length 513;
  Best Local Similarity   99.1%;  Pred. No. 2.6;
  Matches  106;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      82 AAAAAAAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 141

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     142 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188


RESULT   5
CNS0256N
LOCUS       CNS0256N      576 bp    DNA             GSS       12-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            236E22 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL181688
VERSION     AL181688.1  GI:7819767
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 576)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 576)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="236E22"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG236BC11SP1~end :
                     PUC-Ori"
BASE COUNT      554 a      0 c      0 g      5 t      17 others
ORIGIN


  Query Match             96.3%;  Score 105;  DB 13;  Length 576;
  Best Local Similarity   98.1%;  Pred. No. 2.5;
  Matches  105;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||:|||||||||||||||||||||||||||||||||||||||||||||||||||
Db       4 AAAAAAAAAWTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 110


RESULT   6
AL513755/c
LOCUS       AL513755      340 bp    mRNA            EST       13-FEB-2001
DEFINITION  AL513755 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BA008ZH06 3
            prime, mRNA sequence.
ACCESSION   AL513755
VERSION     AL513755.1  GI:12777249
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 340)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .340
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0BA008ZH06"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      176 a      16 c      18 g     121 t       9 others
```

```
ORIGIN


  Query Match             96.0%;  Score 104.6;  DB 10;  Length 340;
  Best Local Similarity   95.4%;  Pred. No. 4;
  Matches  104;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||||| |||| |||||||||||||||||||:|::|||||||||||||||||||||||||
Db    114 AAAATAATAATTAAAAAAAAAAAAAAAAAAAAWAWWAAAAAAAAAAAAAAAAAAAAAAAAA 55

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     54 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 6


RESULT   7
AL514473
LOCUS       AL514473      363 bp    mRNA            EST       13-FEB-2001
DEFINITION  AL514473 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BB010ZE11 3
            prime, mRNA sequence.
ACCESSION   AL514473
VERSION     AL514473.1  GI:12777967
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 363)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .363
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0BB010ZE11"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      231 a      9 c      3 g    112 t      8 others
ORIGIN


  Query Match             96.0%;  Score 104.6;  DB 10;  Length 363;
  Best Local Similarity   97.2%;  Pred. No. 3.8;
  Matches  104;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||:|||:||||||||||||||||||||||||||||||||||||||||||||||||
Db    152 AAAAAAAAAWTTTWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 211

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 258


RESULT   8
AW185559
LOCUS       AW185559      200 bp    mRNA            EST       19-NOV-1999
DEFINITION  se81e02.y1 Gm-c1023 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1023-1107 5', mRNA sequence.
ACCESSION   AW185559
VERSION     AW185559.1  GI:6454949
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 200)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: Genome Systems, Inc. 4633 World
            Parkway Circle St. Louis, Missouri 63134 For further information
            call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
            427-3324 or contact: clones@genomesystems.com or
            info@genomesystems.com web site: www.genomesystems.com
            Putative full length read
            vector to vector length is 201.
FEATURES             Location/Qualifiers
     source          1. .200
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1023-1107"
                     /clone_lib="Gm-c1023"
                     /tissue_type="seed coats of greenhouse grown plants"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     cDNA library was constructed from mRNA isolated from seed
                     coats (100-200mgs) of greenhouse grown plants.  The
                     library was prepared using the Life Technologies
                     pSuperScript cDNA library construction kit.  Complimentary
                     DNA was synthesized from mRNA using a poly (dT) sequence
                     with a Not I restrictions site.  Sal I linkers adapers
                     were ligated to the blunt-ended cDNA fragments followed by
                     Not I digestion.  The cDNA fragments were directionally
                     cloned into the Not I-Sal I restriction site of the
                     pSPORT1 vector.  The ligated cDNA fragments were
                     transformed into E.coli Electromax DH10B host cells. This
                     library was constructed by Dr. Lila Vodkin and Dr. Anu
                     Khanna."
BASE COUNT      121 a     22 c     26 g     31 t
ORIGIN


  Query Match             95.6%;  Score 104.2;  DB 10;  Length 200;
  Best Local Similarity   97.2%;  Pred. No. 6.2;
  Matches  106;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db     15 AAAAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 74

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     75 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 123
```

```
RESULT  9
AL514879/c
LOCUS       AL514879      449 bp    mRNA            EST       13-FEB-2001
DEFINITION  AL514879 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BB014ZE01 3
            prime, mRNA sequence.
ACCESSION   AL514879
VERSION     AL514879.1  GI:12778372
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 449)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .449
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0BB014ZE01"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT       96 a    107 c     19 g    194 t     33 others
ORIGIN


  Query Match             95.6%;  Score 104.2;  DB 10;  Length 449;
  Best Local Similarity   96.3%;  Pred. No. 3.5;
  Matches  103;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||:|||||:|: ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 AAAWTAAAAWTWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 61

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 14


RESULT  10
BG115502
LOCUS       BG115502     1042 bp    mRNA            EST       30-JAN-2001
DEFINITION  602316663F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4416974 5',
            mRNA sequence.
ACCESSION   BG115502
VERSION     BG115502.1  GI:12609008
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1042)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10148  row: n  column: 15
            High quality sequence stop: 391.
FEATURES             Location/Qualifiers
     source          1. .1042
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4416974"
                     /clone_lib="NIH_MGC_88"
                     /tissue_type="duodenal adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
BASE COUNT      446 a    219 c    265 g    112 t
ORIGIN


  Query Match             95.6%;  Score 104.2;  DB 11;  Length 1042;
  Best Local Similarity   97.2%;  Pred. No. 1.9;
  Matches  106;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    253 AAAATAAACATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 312

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          |||||||||||||||||||||||||||||||||||||||||||||||||||| |
Db    313 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 361


RESULT  11
BF984813/c
LOCUS       BF984813     1435 bp    mRNA            EST       23-JAN-2001
DEFINITION  602308063F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4399310 5',
            mRNA sequence.
ACCESSION   BF984813
VERSION     BF984813.1  GI:12387625
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1435)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10102  row: n  column: 15
            High quality sequence start: 8
            High quality sequence stop: 133.
FEATURES             Location/Qualifiers
     source          1. .1435
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4399310"
                     /clone_lib="NIH_MGC_88"
                     /tissue_type="duodenal adenocarcinoma, cell line"
```

```
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: small intestine; Vector: pCMV-SPORT6;
                     Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
                     oligo-dT primed. Average insert size 1.767 kb. Library
                     enriched for full-length clones and constructed by Life
                     Technologies. Note: this is a NIH_MGC Library."
BASE COUNT      589 a    164 c    500 g    182 t
ORIGIN


  Query Match             95.6%;  Score 104.2;  DB 11;  Length 1435;
  Best Local Similarity   97.2%;  Pred. No. 1.5;
  Matches  106;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||||||||  |||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    134 AAAATAAAAAAATAAACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 75

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     74 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 26


RESULT  12
CNS03F8P
LOCUS       CNS03F8P      134 bp    DNA             GSS       17-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            021P15 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL241378
VERSION     AL241378.1  GI:7962147
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 134)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 134)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 134)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .134
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="021P15"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG021CH08SP1-end :
                     PUC-Ori"
BASE COUNT      120 a      2 c      1 g      10 t      1 others
ORIGIN


  Query Match             95.2%;  Score 103.8;  DB 13;  Length 134;
  Best Local Similarity   98.1%;  Pred. No. 8.9;
  Matches  105;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||||||||||  || ||||||||||||||||||||||||||||||||||||||||||||||
Db     28 AAAATAAAAATATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 87

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     88 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 134


RESULT  13
AA117240
LOCUS       AA117240      142 bp    mRNA            EST       15-NOV-1996
DEFINITION  mn19f12.r1 Beddington mouse embryonic region Mus musculus cDNA
            clone IMAGE:538415 5', mRNA sequence.
ACCESSION   AA117240
VERSION     AA117240.1  GI:1672344
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 142)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:325351
            Putative full length read
            vector to vector length is 722
            Seq primer: -40m13 ET
            High quality sequence stop: 106.
FEATURES             Location/Qualifiers
     source          1. .142
                     /organism="Mus musculus"
                     /strain="C57BL6 x DBA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:538415"
                     /clone_lib="Beddington mouse embryonic region"
                     /sex="pooled"
                     /tissue_type="embryo"
                     /dev_stage="7.5dpc"
                     /lab_host="DH12S"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT. Gastrulating embryos were collected at 7.5dpc
                     from C57BL6 x DBA matings, excluding embryos that had
                     developed head folds and all extraembryonic tissues.
                     Average insert size: 1.3 kb (range: 0.5 - 3.0 kb).
                     Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT      135 a      1 c      0 g      6 t
ORIGIN


  Query Match             95.2%;  Score 103.8;  DB 10;  Length 142;
  Best Local Similarity   98.1%;  Pred. No. 8.5;
  Matches  105;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
```

```
          ||||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2 AAAATAAATCTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 61

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 108


RESULT  14
AI340603
LOCUS       AI340603      147 bp    mRNA            EST       29-DEC-1998
DEFINITION  tb31g08.x1 NCI_CGAP_HSC2 Homo sapiens cDNA clone IMAGE:2055998 3'
            similar to contains element OFR repetitive element ;, mRNA
            sequence.
ACCESSION   AI340603
VERSION     AI340603.1  GI:4077530
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 147)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Herbert Morse, M.D., Michael R. Emmert-Buck,
            M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 108.
FEATURES             Location/Qualifiers
     source          1. .147
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2055998"
                     /clone_lib="NCI_CGAP_HSC2"
                     /tissue_type="stem cell 34+/38+"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: bone marrow; Vector: pAMP1; mRNA made from
                     bone marrow, stem cells 34+/38+, cDNA made by oligo-dT
                     priming.  Directionally cloned.  Size-selected on agarose
                     gel, average insert size 400 bp.  Primary library,
                     non-amplified.  "
BASE COUNT      125 a      5 c      9 g      8 t
ORIGIN


  Query Match            95.2%;  Score 103.8;  DB 10;  Length 147;
  Best Local Similarity   98.1%;  Pred. No. 8.3;
  Matches  105;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db     40 AAAAAAAAAATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 99

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 146


RESULT  15
AW087445
LOCUS       AW087445      267 bp    mRNA            EST       15-OCT-1999
DEFINITION  xb98g05.x1 NCI_CGAP_Co21 Homo sapiens cDNA clone IMAGE:2582840 3'
            similar to contains element MER11 repetitive element ;, mRNA
            sequence.
ACCESSION   AW087445
VERSION     AW087445.1  GI:6043250
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 267)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 250.
FEATURES             Location/Qualifiers
     source          1. .267
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2582840"
                     /clone_lib="NCI_CGAP_Co21"
                     /tissue_type="moderately differentiated adenocarcinoma"
                     /lab_host="DH10B"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Normalized to Cot >500.   Average insert size 1.04kb.
                     Normalized version of NCI_CGAP_Co18.   Library constructed
                     by Life Technologies."
BASE COUNT      159 a     27 c     22 g     59 t
ORIGIN


  Query Match            95.2%;  Score 103.8;  DB 10;  Length 267;
  Best Local Similarity   98.1%;  Pred. No. 5.4;
  Matches  105;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||| | ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    129 AAAATTACAATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 188

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    189 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 235


Search completed: January  3, 2002, 09:42:30
Job time: 2733 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 100.12 Seconds
(without alignments)
933.364 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	103.8	95.2	1860	22	AAF72806	Secreted protein g
	2	103.4	94.9	3060	22	AAD02924	Human PRO5723 cDNA
	3	103.4	94.9	3060	22	AAF44263	Human PRO5723 nucl
	4	103.4	94.9	3060	22	AAC91489	Human PRO5723 cDNA
	5	102.6	94.1	225	21	AAC98734	Human colon cancer
c	6	102.6	94.1	396	22	AAF94842	Human ovarian canc
	7	102.6	94.1	464	22	AAH33241	Human colon cancer
	8	102.6	94.1	569	21	AAZ97052	Human secreted pro
	9	102.6	94.1	612	21	AAC99862	Human secreted pro
	10	102.6	94.1	938	22	AAD07722	Human secreted pro
	11	102.6	94.1	1048	21	AAA26373	Human secreted pro
	12	102.6	94.1	1051	21	AAZ91918	Murine mahogany pr
	13	102.6	94.1	1126	21	AAC59409	Human secreted pro
	14	102.6	94.1	1164	22	AAF98699	Human ovarian canc
	15	102.6	94.1	1254	21	AAC99826	Human secreted pro
	16	102.6	94.1	1492	21	AAC98102	Human colon cancer
	17	102.6	94.1	1493	22	AAH34433	Human colon cancer
	18	102.6	94.1	1510	21	AAA39961	Murine TANGO 183 c
	19	102.6	94.1	1733	20	AAV81394	Human tumour antig
	20	102.6	94.1	1868	22	AAD08428	Human secreted pro
c	21	102.6	94.1	9589	14	AAQ38218	NANBH virus strain
	22	102.2	93.8	671	22	AAD05318	Human secreted pro
	23	102.2	93.8	903	21	AAC79727	Human secreted pro
	24	102.2	93.8	2772	22	AAF92074	Human PRO1277 cDNA
	25	102.2	93.8	2773	21	AAZ65009	Membrane-bound pro
	26	102.2	93.8	2773	22	AAF44155	Human PRO1277 (UNQ
	27	101.6	93.2	160	22	AAF24360	Retroviral recombi
	28	101.6	93.2	174	21	AAC98740	Human colon cancer
	29	101.6	93.2	178	22	AAF24361	Retroviral recombi
	30	101.6	93.2	208	21	AAC98557	Human colon cancer
	31	101.6	93.2	208	22	AAF24363	Retroviral recombi
	32	101.6	93.2	712	22	AAH34312	Human colon cancer
	33	101.6	93.2	1708	22	AAF26570	DNA encoding human
	34	101.6	93.2	2240	21	AAC93391	Human secreted pro
	35	101.6	93.2	2240	22	AAD08375	Human secreted pro
	36	101.6	93.2	2240	22	AAS02536	Human secreted pro
	37	101.6	93.2	2427	11	AAQ04107	Human pro-urokinas
	38	101.6	93.2	2510	22	AAH33295	Human colon cancer
	39	101.6	93.2	3899	19	AAV63192	cDNA from clone dt
	40	101.6	93.2	4237	19	AAV61487	Human secreted pro
	41	101.6	93.2	6644	20	AAX33181	Base sequence of t
	42	101.6	93.2	7372	20	AAX33182	Base sequence of t
	43	101.6	93.2	7797	20	AAX33180	Cowpox virus bsr f
	44	101.6	93.2	7996	20	AAX33184	Base sequence of t
c	45	101.4	93.0	396	22	AAF94862	Human ovarian canc

ALIGNMENTS

RESULT 1
AAF72806
ID AAF72806 standard; DNA; 1860 BP.
XX
AC AAF72806;
XX
DT 24-APR-2001 (first entry)
XX
DE Secreted protein gene #8.
XX
KW Secreted protein; human; autoimmune; hyperproliferation;
KW cardiovascular; cerebrovascular; infection; food; ds.
XX
OS Homo sapiens.
XX
PN WO200107459-A1.
XX
PD 01-FEB-2001.
XX
PF 20-JUL-2000; 2000WO-US19735.
XX
PR 23-JUL-1999; 99US-0145220.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Ebner R, Duan RD, Ni J, Soppet DR, Moore PA;
PI Shi Y, Lafleur DW, Olsen HS, Birse CE, Komatsoulis GA;
XX
DR WPI; 2001-123261/13.
XX
PT New isolated nucleic acid encoding 29 secreted proteins, for
PT diagnosing, preventing and treating e.g. autoimmune,
PT hyperproliferative, cardiovascular, and ocular diseases or disorders
PT and microorganism infections -

```
XX
PS   Claim 1; Page 499; 601pp; English.
XX
CC   The present invention relates to 29 human secreted proteins. The
CC   invention is used to prevent autoimmune diseases e.g. rheumatoid
CC   arthritis, hyperproliferative disorders e.g. neoplasms of the
CC   breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC   cerebrovascular disorders e.g. cerebral ischemia, angiogenesis,
CC   nervous system disorders e.g. Alzheimer's disease, infections
CC   caused by bacteria, viruses and fungi and ocular disorders e.g.
CC   corneal infection. Also used in food preparations.
XX
SQ   Sequence 1860 BP; 588 A; 370 C; 410 G; 490 T; 2 other;


  Query Match             95.2%;  Score 103.8;  DB 22;  Length 1860;
  Best Local Similarity   98.1%;  Pred. No. 0.00032;
  Matches  105;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1737 aaaaaaaaaaattaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1796

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1797 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1843


RESULT   2
AAD02924
ID   AAD02924 standard; cDNA; 3060 BP.
XX
AC   AAD02924;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human PRO5723 cDNA (DNA82361).
XX
KW   Human; PRO5723; antiinflammatory; dermatological; immunosuppressive;
KW   antirheumatic; antiarthritic; osteopathic; antianaemic; haemostatic;
KW   antithyroid; antidiabetic; antiviral; antipsoriatic; antiallergic;
KW   antiasthmatic; inhibitor; therapy; systemic lupus erythematosis;
KW   spondyloarthropathy; systemic sclerosis; systemic vasculitis;
KW   sarcoidosis; idiopathic inflammatory myopathy; Sjogren's syndrome;
KW   autoimmune thrombocytopenia; immune-mediated renal disease; hepatitis;
KW   demyelinating polyneuropathy; Guillian-Barre syndrome; Whipple's disease;
KW   hepatobiliary disease; primary biliary cirrhosis; sclerosing cholangitis;
KW   inflammatory bowel disease; gluten-sensitive enteropathy; skin disease;
KW   allergic rhinitis; atopic dermatitis; food hypersensitivity; urticaria;
KW   eosinophilic pneumonia; hypersensitivity pneumonitis; graft rejection;
KW   idiopathic pulmonary fibrosis; graft-versus-host-disease; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             77..1135
FT                   /*tag= a
FT                   /product= "Human PRO5723 protein"
FT   sig_peptide     77..133
FT                   /*tag= b
FT   mat_peptide     134..1132
FT                   /*tag= c
FT                   /product= "Mature human PRO5723 protein"
XX
PN   WO200116319-A2.
XX
PD   08-MAR-2001.
XX
PF   23-AUG-2000; 2000WO-US23522.
XX
PR   31-AUG-1999;   99US-0151733.
PR   01-SEP-1999;   99WO-US20111.
PR   16-DEC-1999;   99WO-US30095.
PR   18-FEB-2000; 2000WO-US04342.
PR   01-MAR-2000; 2000WO-US05601.
PR   30-MAR-2000; 2000WO-US08439.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   05-JUN-2000; 2000US-0209832.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Goddard A,  Godowski PJ,  Gurney AL,  Hillan KJ,  Tumas D;
PI   Watanabe CK,  Wood WI;
XX
DR   WPI; 2001-226690/23.
DR   P-PSDB; AAY72878.
XX
PT   New PRO polypeptides for treating immune related and inflammatory
PT   diseases such as rheumatoid arthritis, systemic vasculitis, asthma,
PT   autoimmune hemolytic anemia, and diabetes mellitus
XX
PS   Claim 2; Fig 9; 118pp; English.
XX
CC   The present sequence is a cDNA (DNA82361 clone) encoding PRO5723
CC   protein. PRO protein, its agonist or antagonist or its antibody which
CC   are capable of enhancing or inhibiting the proliferation of T-lymphocytes
CC   or of increasing the infiltration of inflammatory cells into a tissue
CC   are useful in the diagnosis and treatment of immune-related diseases
CC   in mammals. The PRO protein is useful for treating systemic lupus
CC   erythematosis, rheumatoid arthritis, osteoarthritis, juvenile chronic
CC   arthritis, spondyloarthropathy, systemic sclerosis, idiopathic
CC   inflammatory myopathy, Sjogren's syndrome, systemic vasculitis,
CC   sarcoidosis, autoimmune haemolytic anaemia, autoimmune thrombocytopenia,
CC   thyroiditis, diabetes mellitus, immune-mediated renal disease,
CC   demyelinating disease of the central or peripheral nervous system,
CC   idiopathic demyelinating polyneuropathy, Guillian-Barre syndrome,
CC   chronic inflammatory demyelinating polyneuropathy, hepatobiliary
CC   disease, infectious or autoimmune chronic active hepatitis, primary
CC   biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC   inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC   disease, autoimmune or immune-mediated skin diseases such as bullous
CC   skin disease, erythema multiforme and contact dermatitis, psoriasis,
CC   allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
CC   food hypersensitivity and urticaria, immunologic diseases of the lung
CC   such as eosinophilic pneumonias, idiopathic pulmonary fibrosis, hyper-
CC   sensitivity pneumonitis, transplantation associated diseases such as
CC   graft rejection or graft-versus-host-disease.
XX
SQ   Sequence 3060 BP; 1094 A; 559 C; 569 G; 837 T; 1 other;


  Query Match             94.9%;  Score 103.4;  DB 22;  Length 3060;
  Best Local Similarity   97.2%;  Pred. No. 0.00029;
  Matches  104;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||||||| || |||||||||||:||||||||||||||||||||||||||||||||||||
Db   2929 aaaataaaagttaaaaaaaaaaaaaraaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2988

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2989 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3035


RESULT   3
AAF44263
ID   AAF44263 standard; cDNA; 3060 BP.
XX
AC   AAF44263;
XX
DT   02-APR-2001  (first entry)
XX
DE   Human PRO5723 nucleotide sequence SEQ ID NO:504.
XX
```

KW Human; secreted and transmembrane protein; PRO; cytostatic;
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;
KW diagnostic assay; ss.
XX
OS Homo sapiens.
XX
PN WO200073454-A1.
XX
PD 07-DEC-2000.
XX
PF 30-MAR-2000; 2000WO-US08439.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 23-JUN-1999; 99US-0141037.
PR 07-JUL-1999; 99US-0143048.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 17-AUG-1999; 99US-0149396.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 08-OCT-1999; 99US-0158663.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi CJ, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;
PI Zhang Z;
XX
DR WPI; 2001-032160/04.
DR P-PSDB; AAB65294.
XX
PT PRO polynucleotides used to produce polypeptides used to target
PT bioactive molecules such as toxins, radiolabels or antibodies, to
PT specific cells, to cause targeted cell death -
XX
PS Claim 2; Fig 317; 935pp; English.
XX
CC The present invention describes human secreted and transmembrane PRO
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins
CC can be used for targeted delivery of bioactive molecules, such as
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide
CC sequences, and their fragments, can be used as hybridisation probes, in
CC chromosomal and gene mapping, and in the generation of anti-sense RNA
CC and DNA. They may also be used to produce transgenic animals which are
CC used to develop and screen therapeutically useful reagents. The PRO
CC nucleotide and protein sequence can be used for tissue typing and in
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein
CC sequences given in the exemplification of the present invention.
XX
SQ Sequence 3060 BP; 1094 A; 559 C; 569 G; 837 T; 1 other;

Query Match 94.9%; Score 103.4; DB 22; Length 3060;
Best Local Similarity 97.2%; Pred. No. 0.00029;
Matches 104; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           ||||||||| || ||||||||||||:||||||||||||||||||||||||||||||||||
Db    2929 aaaataaaagttaaaaaaaaaaaaaraaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2988

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    2989 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3035
```

RESULT 4
AAC91489
ID AAC91489 standard; cDNA; 3060 BP.
XX
AC AAC91489;
XX
DT 21-MAR-2001 (first entry)
XX
DE Human PRO5723 cDNA.
XX
KW Human; PRO; antiinflammatory; dermatological; antiarthritic;
KW antirheumatic; cardiant; antianaemic; immunosuppressive; antithyroid;
KW antidiabetic; nootropic; neuroprotective; hepatotropic; virucide;
KW antiallergic; antiasthmatic; immune related disorder;
KW hepatobiliary disease; autoimmune disease; allergy; ss.
XX
OS Homo sapiens.
XX
PN WO200073452-A2.
XX
PD 07-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US15264.
XX
PR 02-JUN-1999; 99WO-US12252.
PR 20-JUL-1999; 99US-0144732.
PR 20-JUL-1999; 99US-0144758.
PR 28-JUL-1999; 99US-0146222.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 29-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28634.
PR 09-DEC-1999; 99US-0170262.
PR 20-DEC-1999; 99WO-US30911.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Baker KP, Chan B, Goddard A, Godowski PJ, Gurney AL;
PI Hebert C, Henzel W, Kabakoff RC, Shelton DL, Tumas D, Watanabe CK;
PI Wood WI;
XX
DR WPI; 2001-025253/03.
DR P-PSDB; AAB50930.
XX
PT Thirty three nucleic acids encoding PRO polypeptides which are useful
PT in the diagnosis and treatment of immune related disorders, e.g.

PT systemic lupus erythematosus, rheumatoid arthritis, osteoarthritis,
PT thyroiditis and diabetes mellitus -
XX
PS Claim 48; Fig 57; 218pp; English.
XX
CC The present sequence is one of thirty three nucleic acids encoding PRO
CC polypeptides. The PRO polypeptides, anti-PRO antibodies, agonists and
CC antagonists are useful for treating and diagnosing immune related
CC disorders such as systemic lupus erythematosus, rheumatoid arthritis,
CC osteoarthritis, juvenile chronic arthritis, spondyloarthropathies,
CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
CC anaemia, autoimmune thrombocytopaenia, thyroiditis, diabetes mellitus,
CC immune-mediated renal disease, demyelinating diseases of the central
CC and peripheral nervous systems (such as multiple sclerosis, idiopathic
CC demyelinating polyneuropathy or Guillain-Barre syndrome, and chronic
CC inflammatory demyelinating polyneuropathy), hepatobiliary diseases
CC (such as infectious, autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis and sclerosing cholangitis),
CC inflammatory bowel disease, gluten-sensitive enteropathy and Whipple's
CC disease, autoimmune or immune-mediated skin diseases (such as bullous
CC skin diseases, erythema multiforme, contact dermatitis, psoriasis),
CC allergic diseases such as asthma, allergic rhinitis, atopic dermatitis,
CC food hypersensitivity and urticaria), immunological diseases of the
CC lung (such as eosinophilic pneumonias, idiopathic pulmonary fibrosis
CC and hypersensitivity pneumonitis), transplantation associated diseases
CC including graft rejection and graft-versus-host diseases.
XX
SQ Sequence 3060 BP; 1094 A; 559 C; 569 G; 837 T; 1 other;

Query Match 94.9%; Score 103.4; DB 22; Length 3060;
Best Local Similarity 97.2%; Pred. No. 0.00029;
Matches 104; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
Db 2929 aaaataaaagttaaaaaaaaaaaaaraaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2988

Qy 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
Db 2989 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 3035

RESULT 5
AAC98734
ID AAC98734 standard; cDNA; 225 BP.
XX
AC AAC98734;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human colon cancer antigen nucleotide sequence SEQ ID NO:744.
XX
KW Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW immunomodulatory; muscular; gynaecological; gastrointestinal;
KW nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW neural disorder; immune system disorder; muscular disorder;
KW reproductive disorder; gastrointestinal disorder; renal disorder;
KW infectious disease; cardiovascular disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200055351-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US05883.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2000-587534/55.
DR P-PSDB; AAB53977.
XX
PT Colon cancer associated gene sequences, referred to as colon cancer
PT antigens, useful for the treatment, prevention, and diagnosis of colon
PT disorders such as colon cancer -
XX
PS Claim 1; Page 1304; 2104pp; English.
XX
CC AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC can be used in gene therapy. The colon cancer antigen polynucleotides,
CC proteins and antibodies to the proteins are useful for the prevention,
CC treatment and diagnosis of colon disorders, such as colon cancer. The
CC polynucleotides may be used in diagnostics and research, such as for
CC chromosome identification, and as hybridisation probes. The proteins
CC may also be used to prevent diseases such as neural disorders, immune
CC system disorders, muscular disorders, reproductive disorders,
CC gastrointestinal disorders, wounds, renal disorders, infectious
CC diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC AAB54007 represent sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 225 BP; 161 A; 11 C; 27 G; 20 T; 6 other;

Query Match 94.1%; Score 102.6; DB 21; Length 225;
Best Local Similarity 96.3%; Pred. No. 0.00094;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
Db 97 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 156

Qy 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
Db 157 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 205

RESULT 6
AAF94842/c
ID AAF94842 standard; cDNA; 396 BP.
XX
AC AAF94842;
XX
DT 23-MAY-2001 (first entry)
XX
DE Human ovarian cancer associated coding sequence SEQ ID NO: 33.
XX
KW Human, ovarian cancer; vaccine; gene therapy; carcinoma; ss.
XX
OS Homo sapiens.
XX
PN WO200118046-A2.
XX
PD 15-MAR-2001.
XX
PF 08-SEP-2000; 2000WO-US24827.
XX
PR 10-SEP-1999; 99US-0394374.
PR 01-MAY-2000; 2000US-0561778.
PR 15-AUG-2000; 2000US-0640173.
PR 07-SEP-2000; 2000US-0656668.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Stolk JA;

```
XX
DR   WPI; 2001-211395/21.
XX
PT   Isolated polypeptides associated with ovarian carcinomas, and the
PT   nucleic acids that encode them, useful for the prevention diagnosis and
PT   treatment of ovarian cancers -
XX
PS   Claim 5; Page 126; 189pp; English.
XX
CC   The present invention provides a number of coding sequences and proteins,
CC   the over-expression of which is associated with ovarian carcinoma/cancer.
CC   These can be used in the diagnosis, treatment and prevention of ovarian
CC   cancer, optionally by gene therapy or in the form of a vaccine. The
CC   present sequence is an example of one of these sequences.
XX
SQ   Sequence 396 BP; 60 A; 17 C; 34 G; 162 T; 123 other;


  Query Match             94.1%;  Score 102.6;  DB 22;  Length 396;
  Best Local Similarity   96.3%;  Pred. No. 0.00076;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 50

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 1


RESULT   7
AAH33241
ID   AAH33241 standard; cDNA; 464 BP.
XX
AC   AAH33241;
XX
DT   03-SEP-2001  (first entry)
XX
DE   Human colon cancer antigen encoding cDNA SEQ ID NO:297.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   colorectal carcinoma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200122920-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26524.
XX
PR   29-SEP-1999;   99US-0157137.
PR   03-NOV-1999;   99US-0163280.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Barash SC,  Birse CE,  Rosen CA;
XX
DR   WPI; 2001-235357/24.
DR   P-PSDB; AAG73810.
XX
PT   Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT   useful for preventing, diagnosing and/or treating colorectal cancers -
XX
PS   Claim 1; Page 2420; 9803pp; English.
XX
CC   AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC   cancer-associated nucleic acid molecules (N) and proteins (P), where
CC   the proteins are collectively known as colon cancer antigens. The colon
CC   cancer antigens have cytostatic activity and can be used in gene
CC   therapy and vaccine production. N and P may be used in the prevention,
CC   diagnosis and treatment of diseases associated with inappropriate P
CC   expression. For example, N and P may be used to treat disorders
CC   associated with decreased expression by rectifying mutations or deletions
CC   in a patient's genome that affect the activity of P by expressing
CC   inactive proteins or to supplement the patients own production of P.
CC   Additionally, N may be used to produce the colon cancer-associated Ps,
CC   by inserting the nucleic acids into a host cell and culturing the cell
CC   to express the proteins. N and P can be used in the prevention, diagnosis
CC   and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
CC   and AAB77789 represent sequences used in the exemplification of the
CC   present invention.
CC   N.B. Pages 666 to 682 and page 7053 of the sequence listing were
CC   missing at time of publication, meaning no sequences are present for
CC   SEQ ID NO:1027 to 1052, 7921 and 7922.
XX
SQ   Sequence 464 BP; 195 A; 93 C; 77 G; 93 T; 6 other;


  Query Match             94.1%;  Score 102.6;  DB 22;  Length 464;
  Best Local Similarity   96.3%;  Pred. No. 0.00071;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     345 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 404

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     405 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 453


RESULT   8
AAZ97052
ID   AAZ97052 standard; cDNA; 569 BP.
XX
AC   AAZ97052;
XX
DT   19-APR-2000  (first entry)
XX
DE   Human secreted protein gene 34 cDNA clone HCHPF68, SEQ ID NO:44.
XX
KW   Human; secreted protein; cancer; tumour; developmental abnormality;
KW   foetal deficiency; blood disorder; immune system disorder; inflammation;
KW   autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
KW   schizophrenia; arthritis; asthma; psoriasis; sepsis; skin disorder;
KW   atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
KW   digestive disorder; endocrine disorder; infection; AIDS; leukaemia;
KW   therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO9966041-A1.
XX
PD   23-DEC-1999.
XX
PF   15-JUN-1999;   99WO-US13418.
XX
PR   16-JUN-1998;   98US-0089507.
PR   16-JUN-1998;   98US-0089508.
PR   16-JUN-1998;   98US-0089509.
PR   16-JUN-1998;   98US-0089510.
PR   22-JUN-1998;   98US-0090112.
PR   22-JUN-1998;   98US-0090113.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Ni J,  Rosen CA,  Wei Y,  Young PE,  Florence KA;
PI   Soppet DR,  Brewer LA,  Endress GA,  Carter KC,  Mucenski M,  Ebner R;
PI   Lafleur DW,  Olsen HS,  Shi Y,  Moore PA,  Komatsoulis G;
XX
DR   WPI; 2000-106100/09.
DR   P-PSDB; AAY86248.
XX
PT   New isolated human genes and the secreted polypeptides they encode,
```

PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX
PS Claim 1; Page 337; 586pp; English.
XX
CC AAZ97019 to AAZ97137 represent 94 isolated human secreted protein genes.
CC AAY86215 to AAY86333 are the secreted proteins encoded by the 94 human
CC genes. This sequence represents a fragment of one of the human secreted
CC proteins. The genes and their corresponding secreted polypeptides are
CC useful for preventing, treating or ameliorating medical conditions,
CC e.g., by protein or gene therapy. Also pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new genes. Specific
CC uses are described for each of the 94 genes, based on which tissues they
CC are most highly expressed in, and include developing products for the
CC diagnosis or treatment of cancer, tumours, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, inflammation, allergies, Alzheimer's and cognitive
CC disorders, schizophrenia, arthritis, asthma, psoriasis, sepsis, skin
CC disorders, atherosclerosis, diabetes, cardiovascular disorders, kidney
CC disorders, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
CC The sequences shown in AAY86334 to AAY86585 represent fragments of the
CC secreted proteins.
XX
SQ Sequence 569 BP; 204 A; 133 C; 137 G; 95 T; 0 other;

Query Match 94.1%; Score 102.6; DB 21; Length 569;
Best Local Similarity 96.3%; Pred. No. 0.00066;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
|||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 453 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 512

Qy 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 513 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 561

RESULT 9
AAC99862
ID AAC99862 standard; cDNA; 612 BP.
XX
AC AAC99862;
XX
DT 13-MAR-2001 (first entry)
XX
DE Human secreted protein gene 45 SEQ ID NO:55.
XX
KW Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW fungicide; ophthalmological; gene therapy; pathological condition;
KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;
KW cerebrovascular disorder; angiogenesis; nervous system disorder;
KW Alzheimer's disease; infection; ocular disorder; corneal infection;
KW wound healing; skin aging; food additive; preservative; ss.
XX
OS Homo sapiens.
XX
PN WO200070042-A1.
XX
PD 23-NOV-2000.
XX
PF 11-MAY-2000; 2000WO-US12788.
XX
PR 13-MAY-1999; 99US-0134068.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Moore PA, Young PE, Komatsoulis GA, Birse CE;
PI Duan RD, Florence KA, Soppet DR;
XX
DR WPI; 2000-679828/66.
DR P-PSDB; AAB56121.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX
PS Claim 1; Page 881; 1065pp; English.
XX
CC The polynucleotide sequences given in AAC99818 to AAC99977 encode the
CC human secreted proteins given in AAB56077 to AAB56362. Human secreted
CC proteins have activities based on the tissues and cells the genes are
CC expressed in. Examples of activities include: immunosuppressive;
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
CC vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
CC virucide; fungicide; and opthalmalogical. The human secreted
CC polynucleotides and proteins can be used to prevent, treat or ameliorate
CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
CC dogs, chickens or sheep. They are also used in diagnosing a pathological
CC condition or susceptibility to a pathological condition. Disorders which
CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The proteins can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used
CC in the exemplification of the present invention.
XX
SQ Sequence 612 BP; 199 A; 139 C; 139 G; 135 T; 0 other;

Query Match 94.1%; Score 102.6; DB 21; Length 612;
Best Local Similarity 96.3%; Pred. No. 0.00064;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
|||| ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 496 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 555

Qy 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 556 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 604

RESULT 10
AAD07722
ID AAD07722 standard; cDNA; 938 BP.
XX
AC AAD07722;
XX
DT 10-AUG-2001 (first entry)
XX
DE Human secreted protein-encoding gene 18 cDNA clone HLYDU43, SEQ ID NO:28.
XX
KW Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angiogenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnerary;
KW cell culture; chemotaxis; food additive; gene therapy;
KW binding partner identification; chromosome 2; ss.

```
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             61..387
FT                   /*tag= a
FT                   /product= "Human secreted protein precursor"
FT   sig_peptide     61..138
FT                   /*tag= b
FT   mat_peptide     139..384
FT                   /*tag= c
FT                   /product= "Human mature secreted protein"
XX
PN   WO200134800-A1.
XX
PD   17-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US30674.
XX
PR   12-NOV-1999;   99US-0164750.
PR   30-JUN-2000; 2000US-0215128.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Ruben SM,  Komatsoulis GA,  Ebner R,  Fiscella M,  Wei P;
XX
DR   WPI; 2001-329085/34.
DR   P-PSDB; AAE03309.
XX
PT   New nucleic acid molecules encoding human secreted proteins, used in
PT   preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT   Parkinson's diseases and cancers -
XX
PS   Claim 1; Page 427; 530pp; English.
XX
CC   AAD07705-AAD07759 represent cDNAs corresponding to 19 human secreted
CC   protein genes, and AAE03292-AAE03346 represent the proteins they encode.
CC   AAE03347-AAE03375 represent human secreted protein fragments or variants.
CC   The genes and their secreted proteins are useful for preventing,
CC   treating or ameliorating medical conditions, e.g., by protein or gene
CC   therapy. Pathological conditions can be diagnosed by determining the
CC   amount of the new protein in a sample or by determining the presence of
CC   mutations in the new genes. Specific uses are described for each of the
CC   19 genes, based on the tissues in which they are most highly expressed,
CC   and include developing products for the diagnosis or treatment of
CC   proliferative disorders, cancer, tumours, foetal and developmental
CC   abnormalities, haematopoietic disorders, diseases of the immune system,
CC   AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC   allergies, neurological disorders (e.g., Alzheimer's disease,
CC   Parkinson's disease), cognitive disorders, schizophrenia, asthma,
CC   skin disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC   cardiovascular disorders, angiogenic disorders, kidney disorders,
CC   gastrointestinal disorders, pregnancy-related disorders, endocrine
CC   disorders, and infections. The proteins can also be used to aid wound
CC   healing and epithelial cell proliferation, to prevent skin aging due to
CC   sunburn, to maintain organs before transplantation, for supporting cell
CC   culture of primary tissues, to regenerate tissues, to identify their
CC   cognate ligands or binding partners, and in chemotaxis, and can be used
CC   as a food additive or preservative to modify storage properties.
CC   Antibodies specific for a protein of the invention can be used in
CC   alleviating symptoms associated with the disorders mentioned above, and
CC   in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC   immunosorbent assay (ELISA). The present sequence represents a human
CC   secreted protein-encoding cDNA of the invention.
XX
SQ   Sequence 938 BP; 388 A; 143 C; 152 G; 253 T; 2 other;


  Query Match             94.1%;  Score 102.6;  DB 22;  Length 938;
  Best Local Similarity   96.3%;  Pred. No. 0.00055;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||  ||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 852

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 901


RESULT  11
AAA26373
ID   AAA26373 standard; cDNA; 1048 BP.
XX
AC   AAA26373;
XX
DT   29-JUN-2000  (first entry)
XX
DE   Human secreted protein gene 28 SEQ ID NO:38.
XX
KW   Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW   antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic;
KW   osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
KW   antipsoriatic; cardiant; gene therapy; cancer; neurological disorder;
KW   immune disease; inflammation; blood disorder; tumour; ss.
XX
OS   Homo sapiens.
XX
PN   WO200006698-A1.
XX
PD   10-FEB-2000.
XX
PF   29-JUL-1999;   99WO-US17130.
XX
PR   30-JUL-1998;   98US-0094657.
PR   05-AUG-1998;   98US-0095486.
PR   06-AUG-1998;   98US-0095454.
PR   06-AUG-1998;   98US-0095455.
PR   12-AUG-1998;   98US-0096319.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Komatsoulis GA,  Rosen CA,  Ruben SM,  Duan R,  Moore PA,  Shi Y;
PI   Lafleur D,  Wei Y,  Ni J,  Florence KA,  Young PE,  Brewer LA;
PI   Soppet DR,  Endress GA,  Ebner R,  Olsen HS,  Mucenski M;
XX
DR   WPI; 2000-195282/17.
DR   P-PSDB; AAY91478.
XX
PT   New isolated human genes and the secreted polypeptides they encode,
PT   useful for diagnosis and treatment of e.g. cancers, neurological
PT   disorders, immune diseases, inflammation or blood disorders -
XX
PS   Claim 1; Page 392; 634pp; English.
XX
CC   The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC   human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC   proteins can have activities based on the tissues and cells they are
CC   expressed in. Examples of the activities are: cytostatic;
CC   immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC   antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC   antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC   corresponding secreted proteins are useful for preventing, treating or
CC   ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC   pathological conditions can be diagnosed by determining the amount of the
CC   proteins in a sample or by determining the presence of mutations in the
CC   polynucleotides. Specific uses are described for each of the
CC   polynucleotides, based on which tissues they are most highly expressed
CC   in, and include developing products for the diagnosis or treatment of
CC   cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC   and foetal deficiencies, blood disorders, diseases of the immune system,
CC   autoimmune diseases, hepatic and renal disease, inflammation,
CC   allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC   osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC   transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
```

```
CC   cardiovascular disorders, reproductive disorders, gastrointestinal
CC   disorders, respiratory disorders and metabolic disorders. The proteins
CC   or polynucleotides can also be used as food additives or preservatives.
CC   The proteins are also useful for identifying their binding partners.
CC   AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 1048 BP; 372 A; 203 C; 182 G; 291 T; 0 other;


  Query Match             94.1%;  Score 102.6;  DB 21;  Length 1048;
  Best Local Similarity   96.3%;  Pred. No. 0.00052;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     932 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 991

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     992 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 1040


RESULT  12
AAZ91918
ID   AAZ91918 standard; cDNA; 1051 BP.
XX
AC   AAZ91918;
XX
DT   08-JUN-2000  (first entry)
XX
DE   Murine mahogany protein coding sequence akm1004.
XX
KW   Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;
KW   weight regulation; cell therapy; body weight disorder; cachexia;
KW   anorexia; hyperpigmentation; increased metabolic rate disorder;
KW   hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS   Mus sp.
XX
PN   WO200005373-A2.
XX
PD   03-FEB-2000.
XX
PF   21-JUL-1999;   99WO-US16484.
XX
PR   21-JUL-1998;   98US-0093630.
PR   20-OCT-1998;   98US-0104978.
PR   05-FEB-1999;   99US-0245041.
XX
PA   (MILL-) MILLENIUM PHARM INC.
XX
PI   Moore K,  Nagle DL;
XX
DR   WPI; 2000-195103/17.
DR   P-PSDB; AAY81805.
XX
PT   New human and murine mahogany genes, useful, e.g. for diagnosis and
PT   treatment of body weight disorders -
XX
PS   Claim 1; Fig 9a; 188pp; English.
XX
CC   This sequence represents a murine mahogany gene of the invention.
CC   The mahogany genes are used: (i) to produce recombinant mahogany (mg)
CC   proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
CC   therapeutics; (iii) as a source of diagnostic probes and primers for
CC   detecting expression of mg genes or mutations, regulatory defects, in
CC   this gene, or for isolation of related sequences; and
CC   (iv) in (cell-based) gene therapy. (II) are used to raise specific
CC   antibodies (Ab); to identify other (extra)cellular products involved in
CC   weight regulation, and to screen for agents that disrupt interaction
CC   between (II) and other macromolecules. The Ab are used to detect abnormal
CC   levels (or function) of (II) (for diagnosis, prognosis or monitoring of
CC   treatment); to evaluate (II)-expressing cells intended for cell therapy,
CC   and as therapeutic mg inhibitors. Cells that express the mg gene (or
CC   contain the mg polypeptide) are used to identify agents (A) that modulate
CC   mg activity. (A) are potentially useful for the treatment of body weight
CC   disorders, particularly obesity, cachexia or anorexia, or other
CC   conditions associated with the mg gene such as hyperpigmentation,
CC   hyperphagia and disorders that result in increased metabolic rate.
XX
SQ   Sequence 1051 BP; 306 A; 243 C; 313 G; 189 T; 0 other;


  Query Match             94.1%;  Score 102.6;  DB 21;  Length 1051;
  Best Local Similarity   96.3%;  Pred. No. 0.00052;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
Db      32 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 91

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      92 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 140


RESULT  13
AAC59409
ID   AAC59409 standard; cDNA; 1126 BP.
XX
AC   AAC59409;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Human secreted protein cDNA #18.
XX
KW   Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW   antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW   vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW   cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; human; secreted protein; ss.
XX
OS   Homo sapiens.
XX
PN   WO200056765-A1.
XX
PD   28-SEP-2000.
XX
PF   16-MAR-2000; 2000WO-US06823.
XX
PR   19-MAR-1999;   99US-0125364.
PR   08-DEC-1999;   99US-0169623.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis G;
XX
DR   WPI; 2000-602215/57.
DR   P-PSDB; AAB33979.
XX
PT   Nucleic acid molecules encoding human secreted proteins, used in
PT   preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT   Parkinson's diseases and cancers -
XX
PS   Claim 1; Page 337; 410pp; English.
XX
CC   The invention relates to the isolation of genes AAC59392-C59439 encoding
CC   48 human secreted proteins AAB33963-B34006. The genes can be used to
CC   generate fusion proteins by linking to the gene for the human
CC   immunoglobulin G Fc portion (SEQID1) for increasing the stability of
CC   the fusion protein as compared to the human protein only. The genes and
CC   proteins are useful for preventing, ameliorating or treating medical
CC   conditions, e.g. by protein or gene therapy. The genes are isolated
CC   from a range of human tissues disclosed in the specification. The
CC   nucleic acids, proteins, antibodies and (ant)agonists are useful in
```

```
CC   the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC   and ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC   marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC   (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC   haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC   disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC   colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
CC   wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC   epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC   and parasitic infections.
XX
SQ   Sequence 1126 BP; 368 A; 291 C; 143 G; 324 T; 0 other;


  Query Match             94.1%;  Score 102.6;  DB 21;  Length 1126;
  Best Local Similarity   96.3%;  Pred. No. 0.00051;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     986 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1045

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1046 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 1094


RESULT  14
AAF98699
ID   AAF98699 standard; DNA; 1164 BP.
XX
AC   AAF98699;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human ovarian cancer cell expressed sequence 10799.
XX
KW   Human; ovarian cancer; identification; detection; characterisation;
KW   tumour; kinase; marker; cytostatic; antisense gene therapy; ds.
XX
OS   Homo sapiens.
XX
PN   WO200118542-A2.
XX
PD   15-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24199.
XX
PR   03-SEP-1999;   99US-0152547.
PR   16-MAR-2000; 2000US-0190347.
PR   21-MAR-2000; 2000US-0191321.
PR   31-MAY-2000; 2000US-0208382.
PR   20-JUL-2000; 2000US-0220467.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lee J,  Thompsho P,  Lillie J;
XX
DR   WPI; 2001-211428/21.
XX
PT   Detection, assessment, prevention and therapy of ovarian cancer,
PT   comprises detecting changes in the expression of a variety of markers -
XX
PS   Claim 1; Page 1002; 1198pp; English.
XX
CC   The present invention describes a method for assessing whether a patient
CC   is afflicted with ovarian cancer by comparing: (1) the expression of a
CC   marker (I) (see AAF98594 to AAF98730), in a patient sample; and (2) the
CC   normal level of expression of (I) in a control non-ovarian cancer
CC   sample, where a significant difference between the level of expression
CC   in (a) and (b) is an indication that the patient is afflicted with
CC   ovarian cancer. (I) have cytostatic activities and can be used in
CC   antisense gene therapy. The method, compositions and kits from the
CC   present invention can be used for: (1) assessing and treating ovarian
CC   cancer; (2) making isolated hybridoma, which produces an antibody useful
CC   for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a
CC   patient. AAF98573 to AAF98593 represent human kinase marker primers and
CC   probes which are used in the exemplification of the present invention.
XX
SQ   Sequence 1164 BP; 420 A; 176 C; 195 G; 373 T; 0 other;


  Query Match             94.1%;  Score 102.6;  DB 22;  Length 1164;
  Best Local Similarity   96.3%;  Pred. No. 0.0005;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1048 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1107

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1108 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 1156


RESULT  15
AAC99826
ID   AAC99826 standard; cDNA; 1254 BP.
XX
AC   AAC99826;
XX
DT   13-MAR-2001  (first entry)
XX
DE   Human secreted protein gene 9 SEQ ID NO:19.
XX
KW   Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
KW   antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
KW   cerebroprotective; nootropic; neuroprotective; antibacterial; virucide;
KW   fungicide; ophthalmological; gene therapy; pathological condition;
KW   autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
KW   neoplasm; cardiovascular disorder; cardiac arrest; cerebral ischaemia;
KW   cerebrovascular disorder; angiogenesis; nervous system disorder;
KW   Alzheimer's disease; infection; ocular disorder; corneal infection;
KW   wound healing; skin aging; food additive; preservative; ss.
XX
OS   Homo sapiens.
XX
PN   WO200070042-A1.
XX
PD   23-NOV-2000.
XX
PF   11-MAY-2000; 2000WO-US12788.
XX
PR   13-MAY-1999;   99US-0134068.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM,  Moore PA,  Young PE,  Komatsoulis GA,  Birse CE;
PI   Duan RD,  Florence KA,  Soppet DR;
XX
DR   WPI; 2000-679828/66.
DR   P-PSDB; AAB56085.
XX
PT   Isolated nucleic acid molecule encoding a human secreted protein is
PT   used in preventing, treating or ameliorating a medical condition  -
XX
PS   Claim 1; Page 860-861; 1065pp; English.
XX
CC   The polynucleotide sequences given in AAC99818 to AAC99977 encode the
CC   human secreted proteins given in AAB56077 to AAB56362. Human secreted
CC   proteins have activities based on the tissues and cells the genes are
CC   expressed in. Examples of activities include: immunosuppressive;
CC   antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;
CC   vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial;
CC   virucide; fungicide; and opthalmalogical. The human secreted
CC   polynucleotides and proteins can be used to prevent, treat or ameliorate
```

CC a medical condition in e.g. humans, mice, rabbits, goats, horses, cats,
CC dogs, chickens or sheep. They are also used in diagnosing a pathological
CC condition or susceptibility to a pathological condition. Disorders which
CC are diagnosed or treated include autoimmune diseases e.g. rheumatoid
CC arthritis, hyperproliferative disorders e.g. neoplasms of the breast or
CC liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular
CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders
CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and
CC fungi and ocular disorders e.g. corneal infection. The proteins can also
CC be used to aid wound healing and epithelial cell proliferation, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, for supporting cell culture of primary tissues, to
CC regenerate tissues and in chemotaxis. The proteins can also be used as a
CC food additive or preservative to increase or decrease storage
CC capabilities. AAC99809 to AAC99817 and AAB56076 represent sequences used
CC in the exemplification of the present invention.
XX
SQ Sequence 1254 BP; 348 A; 383 C; 323 G; 197 T; 3 other;

Query Match 94.1%; Score 102.6; DB 21; Length 1254;
Best Local Similarity 96.3%; Pred. No. 0.00049;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1197

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 1246
```

Search completed: January 3, 2002, 08:59:01
Job time : 124 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 09:21:52 ; Search time 38.35 Seconds
(without alignments)
643.705 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 351203 seqs, 113238999 residues

Word size : 0

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	97	89.0	144	1	US-08-702-344-26	Sequence 26, Appl
2	97	89.0	1117	4	US-09-247-373B-33	Sequence 33, Appl
3	96	88.1	1051	4	US-09-245-041-10	Sequence 10, Appl
4	96	88.1	1474	4	US-08-821-994-64	Sequence 64, Appl
5	96	88.1	1733	3	US-09-073-569-1	Sequence 1, Appli
c 6	96	88.1	9589	1	US-07-925-695-1	Sequence 1, Appli
c 7	96	88.1	9589	1	US-07-925-695-2	Sequence 2, Appli
8	95	87.2	140	1	US-08-628-417-5	Sequence 5, Appli
9	95	87.2	240	1	US-08-628-417-6	Sequence 6, Appli
10	95	87.2	1066	1	US-08-157-101A-4	Sequence 4, Appli
11	95	87.2	1582	3	US-08-545-196B-10	Sequence 10, Appl
12	95	87.2	1582	3	US-08-545-196B-12	Sequence 12, Appl
13	94	86.2	98	1	US-08-088-658-42	Sequence 42, Appl
14	94	86.2	98	2	US-08-471-907A-42	Sequence 42, Appl
15	94	86.2	111	4	US-09-297-535-23	Sequence 23, Appl
16	94	86.2	117	1	US-08-702-344-3	Sequence 3, Appli
17	94	86.2	121	4	US-09-297-535-20	Sequence 20, Appl
18	94	86.2	1813	4	US-09-071-224-3	Sequence 3, Appli
19	94	86.2	1882	4	US-09-370-253-1	Sequence 1, Appli
20	94	86.2	2082	2	US-08-785-310A-2	Sequence 2, Appli
21	94	86.2	2184	4	US-08-955-918C-1	Sequence 1, Appli
22	94	86.2	2447	2	US-09-014-969-14	Sequence 14, Appl
23	94	86.2	6671	1	US-08-280-443-1	Sequence 1, Appli
24	94	86.2	6671	1	US-08-457-459-1	Sequence 1, Appli
25	94	86.2	6671	1	US-08-555-678-1	Sequence 1, Appli
26	94	86.2	6671	5	PCT-US95-02275-1	Sequence 1, Appli
27	92	84.4	2671	6	5168051-9	Patent No. 5168051
28	91	83.5	1411	4	US-08-964-127-5	Sequence 5, Appli
29	90	82.6	1641	1	US-08-300-903A-8	Sequence 8, Appli
30	89	81.7	1098	3	US-09-248-335-35	Sequence 35, Appl
31	88	80.7	2852	3	US-09-027-137-2	Sequence 2, Appli
32	87	79.8	1198	3	US-09-248-335-27	Sequence 27, Appl
33	87	79.8	1817	1	US-08-473-981A-5	Sequence 5, Appli
34	87	79.8	1817	2	US-08-474-087-5	Sequence 5, Appli
35	86	78.9	2239	4	US-09-196-390-1	Sequence 1, Appli
36	84	77.1	630	1	US-08-185-414E-1	Sequence 1, Appli
37	83	76.1	1602	1	US-08-530-950-3	Sequence 3, Appli
38	83	76.1	1602	3	US-08-888-429A-3	Sequence 3, Appli
39	83	76.1	1602	4	US-09-149-879-3	Sequence 3, Appli
40	80	73.4	635	1	US-08-455-633A-35	Sequence 35, Appl
41	80	73.4	635	1	US-08-416-336-5	Sequence 5, Appli
42	80	73.4	635	2	US-08-456-460C-35	Sequence 35, Appl
43	80	73.4	635	5	PCT-US94-05354-35	Sequence 35, Appl
44	80	73.4	1307	2	US-08-960-022-17	Sequence 17, Appl
45	80	73.4	1736	3	US-09-182-816-22	Sequence 22, Appl

ALIGNMENTS

```
RESULT   1
US-08-702-344-26
; Sequence 26, Application US/08702344
; Patent No. 5723315
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobs, Kenneth
;    APPLICANT:  McCoy, John
;    APPLICANT:  LaVallie, Edward
;    APPLICANT:  Racie, Lisa
;    APPLICANT:  Merberg, David
;    APPLICANT:  Treacy, Maurice
;    APPLICANT:  Spaulding, Vikki
;    TITLE OF INVENTION:  SECRETED PROTEINS AND POLYNUCLEOTIDES
;    TITLE OF INVENTION:  ENCODING THEM
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/702,344
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  144 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-702-344-26


  Query Match             89.0%;  Score 97;  DB 1;  Length 144;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-22;
  Matches   97;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
;      ZIP:  98102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/073,569
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sawislak, Deborah A
;      REGISTRATION NUMBER:  37,438
;      REFERENCE/DOCKET NUMBER:  97-14
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  206-442-6672
;      TELEFAX:  206-442-6678
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1733 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  34...1344
;      OTHER INFORMATION:
US-09-073-569-1


  Query Match             88.1%;  Score 96;  DB 3;  Length 1733;
  Best Local Similarity   100.0%;  Pred. No. 2e-22;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1630 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1689

Qy      74 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           |||||||||||||||||||||||||||||||||
Db    1690 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 1725


RESULT   6
US-07-925-695-1/c
; Sequence 1, Application US/07925695
; Patent No. 5428145
;  GENERAL INFORMATION:
;    APPLICANT:  OKAMOTO, Hiroaki
;    APPLICANT:  NAKAMURA, Tetsuo
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOME,
;    TITLE OF INVENTION:  POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
;    TITLE OF INVENTION:  DETECTION SYSTEMS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Beveridge, DeGrandi, Weilacher & Young
;      STREET:  1850 M Street, N.W., Suite 800
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/925,695
;      FILING DATE:  19920807
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 287402/91
;      FILING DATE:  09-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 360441/91
;      FILING DATE:  05-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weilacher, Robert G.
;      REGISTRATION NUMBER:  20,531
;      REFERENCE/DOCKET NUMBER:  06/87-48009
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 659-2811
;      TELEFAX:  (202) 659-1462
;      TELEX:  WUI 64470
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9589 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-07-925-695-1


  Query Match             88.1%;  Score 96;  DB 1;  Length 9589;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-22;
  Matches   96;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    9575 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9516

Qy      74 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           |||||||||||||||||||||||||||||||||
Db    9515 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480


RESULT   7
US-07-925-695-2/c
; Sequence 2, Application US/07925695
; Patent No. 5428145
;  GENERAL INFORMATION:
;    APPLICANT:  OKAMOTO, Hiroaki
;    APPLICANT:  NAKAMURA, Tetsuo
;    TITLE OF INVENTION:  NON-A, NON-B HEPATITIS VIRUS GENOME,
;    TITLE OF INVENTION:  POLYNUCLEOTIDES, POLYPEPTIDES, ANTIGEN, ANTIBODY AND
;    TITLE OF INVENTION:  DETECTION SYSTEMS
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Beveridge, DeGrandi, Weilacher & Young
;      STREET:  1850 M Street, N.W., Suite 800
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  US
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/925,695
;      FILING DATE:  19920807
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 287402/91
;      FILING DATE:  09-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 360441/91
;      FILING DATE:  05-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weilacher, Robert G.
;      REGISTRATION NUMBER:  20,531
```

```
; Patent No. 5808032
;  GENERAL INFORMATION:
;    APPLICANT:  KURIHARA, TATSUYA
;    APPLICANT:  MATSUKURA, SHIGEKAZU
;    APPLICANT:  TSURUOKA, NOBUO
;    APPLICANT:  ARIMA, KENJI
;    APPLICANT:  NISHIHARA, TATSURO
;    TITLE OF INVENTION:  ANTI-HBs ANTIBODY GENES AND EXPRESSION
;    TITLE OF INVENTION:  PLASMIDS THEREFOR
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PILLSBURY, MADISON & SUTRO
;      STREET:  1100 NEW YORK AVENUE, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/157,101A
;      FILING DATE:  05-APR-1994
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  TITUS, MARLANA K
;      REGISTRATION NUMBER:  35843
;      REFERENCE/DOCKET NUMBER:  9437/204199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-861-3711
;      TELEFAX:  202-822-0944
;      TELEX:  6714627 CUCH
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1066 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-157-101A-4


  Query Match             87.2%;  Score 95;  DB 1;  Length 1066;
  Best Local Similarity   100.0%;  Pred. No. 4.6e-22;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 73
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     947 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1006

Qy      74 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaag 108
           |||||||||||||||||||||||||||||||||||||
Db    1007 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG 1041


RESULT  11
US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
;  GENERAL INFORMATION:
;    APPLICANT:  MELKI, JUDITH
;    APPLICANT:  MUNNICH, ARNOLD
;    TITLE OF INVENTION:  SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
;    TITLE OF INVENTION:  FOR SPINAL MUSCULAR ATROPHY
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH, LLP
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,196B
;      FILING DATE:  19-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FARACI, C. J.
;      REGISTRATION NUMBER:  32,350
;      REFERENCE/DOCKET NUMBER:  2121-110P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1582 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-545-196B-10


  Query Match             87.2%;  Score 95;  DB 3;  Length 1582;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-22;
  Matches   95;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 70
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1488 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1547

Qy      71 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 105
           |||||||||||||||||||||||||||||||||||||
Db    1548 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1582


RESULT  12
US-08-545-196B-12
; Sequence 12, Application US/08545196B
; Patent No. 6080577
;  GENERAL INFORMATION:
;    APPLICANT:  MELKI, JUDITH
;    APPLICANT:  MUNNICH, ARNOLD
;    TITLE OF INVENTION:  SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
;    TITLE OF INVENTION:  FOR SPINAL MUSCULAR ATROPHY
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BIRCH, STEWART, KOLASCH AND BIRCH, LLP
;      STREET:  PO BOX 747
;      CITY:  FALLS CHURCH
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22040-0747
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/545,196B
;      FILING DATE:  19-OCT-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FARACI, C. J.
;      REGISTRATION NUMBER:  32,350
;      REFERENCE/DOCKET NUMBER:  2121-110P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 205-8000
;      TELEFAX:  (703) 205-8050
```

RESULT 15
US-09-297-535-23
; Sequence 23, Application US/09297535
; Patent No. 6268199
; GENERAL INFORMATION:
; APPLICANT: Meulenberg, Johanna J.M.
; APPLICANT: Pol, Johannes M.A.
; APPLICANT: Bos-de Ruijter, Judy N.A.
; TITLE OF INVENTION: Infectious clones of RNA viruses and vaccines and
; TITLE OF INVENTION: diagnostic assays derived thereof.
; FILE REFERENCE: P20749US00
; CURRENT APPLICATION NUMBER: US/09/297,535
; CURRENT FILING DATE: 1999-10-12
; EARLIER APPLICATION NUMBER: PCT/NL97/00593
; EARLIER FILING DATE: 1997-10-29
; EARLIER APPLICATION NUMBER: 96203024.3
; EARLIER FILING DATE: 1996-10-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 111
; TYPE: RNA
; ORGANISM: Lelystad virus
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: (1)..(111)
; OTHER INFORMATION: /note="3' end"
US-09-297-535-23

Query Match 86.2%; Score 94; DB 4; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.8e-21;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 14 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 73
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60

Qy 74 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
||||||||||||||||||||||||||||||||||||
Db 61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 94

Search completed: January 3, 2002, 10:25:57
Job time: 3845 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 3, 2002, 08:56:57 ; Search time 1340.09 Seconds
(without alignments)
1341.845 Million cell updates/sec

Title: US-09-594-596-1
Perfect score: 109
Sequence: 1 aaaataaaaatttaaaaaaa.....aaaaaaaaaaaaaaaaaagg 109

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match	Length	DB	ID	Description
1		105.4	96.7	1801	14	SCU78817	U78817 Saccharomyc
2		103.4	94.9	2506	9	AK026927	AK026927 Homo sapi
3		103.4	94.9	3060	6	AX055476	AX055476 Sequence
4		103.4	94.9	3060	6	AX089948	AX089948 Sequence
5		102.8	94.3	596	9	BC008387	BC008387 Homo sapi
6	c	102.6	94.1	396	6	AX093215	AX093215 Sequence
7		102.6	94.1	1164	6	AX098194	AX098194 Sequence
8		102.6	94.1	1699	9	HSM801509	AL133640 Homo sapi
9	c	102.6	94.1	9589	6	E07361	E07361 gRNA of Hep
10	c	102.6	94.1	9589	6	E07362	E07362 cDNA of Hep
11	c	102.6	94.1	9589	6	I12861	I12861 Sequence 2
12	c	102.6	94.1	9589	14	HPCPOLP	D00944 Hepatitis C
13		102.6	94.1	214139	2	AC022129	AC022129 Homo sapi
14		102.2	93.8	161	6	A08916	A08916 H.sapiens (
15		102.2	93.8	353	11	CNS06EJ3	AL395173 T7 end of
16		102.2	93.8	951	9	HSM801406	AL133560 Homo sapi
17		102.2	93.8	2417	9	HSM800967	AL117457 Homo sapi
18		102.2	93.8	2773	6	AX092302	AX092302 Sequence
19		101.8	93.4	1843	9	BC009470	BC009470 Homo sapi
20		101.6	93.2	837	9	S78214	S78214 APC=tumor s
21		101.6	93.2	872	6	AX068322	AX068322 Sequence
22		101.6	93.2	1243	9	HSM802621	AL162083 Homo sapi
23		101.6	93.2	1489	9	AB048953	AB048953 Macaca fa
24		101.6	93.2	1892	9	HSM801812	AL136844 Homo sapi
25		101.6	93.2	2118	9	HSM800970	AL117460 Homo sapi
26		101.6	93.2	2296	9	HSM802712	AL359601 Homo sapi
27		101.6	93.2	2621	10	MUSBRED	L31396 Mus musculu
28		101.6	93.2	3924	10	MUSBREDA	L31397 Mus musculu
29		101.6	93.2	5379	9	HSM801286	AL122123 Homo sapi
30		101.6	93.2	6644	6	E23356	E23356 Virus vecto
31		101.6	93.2	7372	6	E23357	E23357 Virus vecto
32		101.6	93.2	7797	6	E23355	E23355 Virus vecto
33		101.6	93.2	7996	6	E23359	E23359 Virus vecto
34		101.6	93.2	74896	2	AC093310	AC093310 Homo sapi
35		101.6	93.2	169170	2	AC093281	AC093281 Homo sapi
36	c	101.6	93.2	223538	2	AC087150	AC087150 Mus muscu
37		101.6	93.2	256673	2	AC087146	AC087146 Mus muscu
38	c	101.4	93.0	396	6	AX093235	AX093235 Sequence
39		101.2	92.8	381	3	AF159974	AF159974 Buthus ma
40		101.2	92.8	2785	9	HSM800418	AL050108 Homo sapi
41	c	101.2	92.8	45685	2	AC087168	AC087168 Homo sapi
42		101	92.7	1569	9	HSM801353	AL133080 Homo sapi
43	c	101	92.7	9580	14	AF054250	AF054250 Hepatitis
44		101	92.7	154272	2	AP001888	AP001888 Homo sapi
45		100.6	92.3	117	6	I89931	I89931 Sequence 3

ALIGNMENTS

RESULT 1
SCU78817
LOCUS SCU78817 1801 bp RNA VRL 29-JUL-1997
DEFINITION Saccharomyces cerevisiae killer virus M1, complete genome.
ACCESSION U78817
VERSION U78817.1 GI:1699029
KEYWORDS .
SOURCE Saccharomyces cerevisiae killer virus M1.
ORGANISM Saccharomyces cerevisiae killer virus M1
Viruses; unclassified viruses.
REFERENCE 1 (bases 1 to 1801)
AUTHORS Russell,P.J., Bennett,A.M., Love,Z. and Baggott,D.M.
TITLE Cloning, sequencing and expression of a full-length cDNA copy of the M1 double-stranded RNA virus from the yeast, Saccharomyces cerevisiae
JOURNAL Yeast 13 (9), 829-836 (1997)
MEDLINE 97377990
REFERENCE 2 (bases 1 to 1801)
AUTHORS Russell,P.J., Bennett,A.M., Love,Z. and Baggott,D.M.
TITLE Direct Submission

```
DEFINITION  Sequence 9 from Patent WO0116319.
ACCESSION   AX089948
VERSION     AX089948.1  GI:13443985
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3060)
  AUTHORS   Goddard,A., Godowski,P.J., Gurney,A.L., Hillan,K.J., Tumas,D.,
            Watanabe,C.K. and Wood,W.I.
  TITLE     Compositions and methods for the treatment of immune related
            diseases
  JOURNAL   Patent: WO 0116319-A 9 08-MAR-2001;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .3060
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1094 a    559 c    569 g    837 t      1 others
ORIGIN


  Query Match            94.9%;  Score 103.4;  DB 6;  Length 3060;
  Best Local Similarity  97.2%;  Pred. No. 0.006;
  Matches  104;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           ||||||||| || ||||||||||||:|||||||||||||||||||||||||||||||||||||
Db    2929 AAAATAAAAGTTAAAAAAAAAAAAAARAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2988

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    2989 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3035


RESULT   5
BC
LOCUS       BC008387      596 bp    mRNA            PRI       12-JUL-2001
DEFINITION  Homo sapiens, clone IMAGE:4052341, mRNA.
ACCESSION   BC008387
VERSION     BC008387.1  GI:14249997
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 596)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-MAY-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 16 Row: p Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Similarity but not
            identity to protein.

FEATURES             Location/Qualifiers
     source          1. .596
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4052341"
                     /tissue_type="Kidney, hypernephroma"
                     /clone_lib="NIH_MGC_58"
                     /lab_host="DH10B"
                     /note="Vector: pDNR-LIB"
BASE COUNT     286 a     82 c    107 g    121 t
ORIGIN


  Query Match            94.3%;  Score 102.8;  DB 9;  Length 596;
  Best Local Similarity  98.1%;  Pred. No. 0.009;
  Matches  104;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 aaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 61
           ||||||||| | ||||||||||||||||||||||||||||||||||||||||||||||||
Db     441 AAATAAAAAATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 500

Qy      62 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     501 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 546


RESULT   6
AX093215/c
LOCUS       AX093215      396 bp    DNA             PAT       30-MAR-2001
DEFINITION  Sequence 33 from Patent WO0118046.
ACCESSION   AX093215
VERSION     AX093215.1  GI:13509664
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Xu,J. and Stolk,J.A.
  TITLE     Ovarian tumor sequences and methods of use therefor
  JOURNAL   Patent: WO 0118046-A 33 15-MAR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    1. .396
                     /note="n = A,T,C or G"
BASE COUNT      60 a     17 c     34 g    162 t    123 others
ORIGIN


  Query Match            94.1%;  Score 102.6;  DB 6;  Length 396;
  Best Local Similarity  96.3%;  Pred. No. 0.01;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 50

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db      49 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 1


RESULT   7
AX098194
LOCUS       AX098194      1164 bp    DNA             PAT       30-MAR-2001
DEFINITION  Sequence 106 from Patent WO0118542.
ACCESSION   AX098194
VERSION     AX098194.1  GI:13515278
KEYWORDS    .
SOURCE      human.
```

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480
```

RESULT 10
E07362/c
LOCUS E07362 9589 bp RNA PAT 29-SEP-1997
DEFINITION cDNA of Hepatitis non-A non-B virus.
ACCESSION E07362
VERSION E07362.1 GI:2175501
KEYWORDS JP 1994121689-A/2.
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 9589)
AUTHORS Okamoto,H. and Nakamura,T. .
TITLE NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE, ANTIGEN AND ANTIBODY DETECTION SYSTEM
JOURNAL Patent: JP 1994121689-A 2 06-MAY-1994;
NAKAMURA TETSUO
COMMENT OS Hepatitis non-A non-B virus
PN JP 1994121689-A/2
PD 06-MAY-1994
PF 09-AUG-1991 JP 1991287402
PI OKAMOTO HIROAKI, NAKAMURA TETSUO
PC C12N15/51,A61K39/29,A61K39/395,A61K39/395,C07K13/00,C12P21/02,
PC C12P21/08,
PC G01N33/53,G01N33/576,G01N33/577//A61B10/00;
CC strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FH
FT source 1. .9589
FT /organism='Hepatitis non-A non-B virus' FT
/strain='HCV-J6'.
FEATURES Location/Qualifiers
source 1. .9589
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 1968 a 2820 c 2635 g 2166 t
ORIGIN

Query Match 94.1%; Score 102.6; DB 6; Length 9589;
Best Local Similarity 96.3%; Pred. No. 0.0061;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480
```

RESULT 11
I12861/c
LOCUS I12861 9589 bp DNA PAT 26-JUL-1995
DEFINITION Sequence 2 from patent US 5428145.
ACCESSION I12861
VERSION I12861.1 GI:910242
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 9589)
AUTHORS Okamoto,H. and Nakamura,T.
TITLE Non-A, non-B, hepatitis virus genome, polynucleotides, polypeptides, antigen, antibody and detection systems
JOURNAL Patent: US 5428145-A 2 27-JUN-1995;
FEATURES Location/Qualifiers
source 1. .9589
/organism="unknown"
BASE COUNT 1968 a 2820 c 2635 g 2166 t
ORIGIN

Query Match 94.1%; Score 102.6; DB 6; Length 9589;
Best Local Similarity 96.3%; Pred. No. 0.0061;
Matches 105; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480
```

RESULT 12
HPCPOLP/c
LOCUS HPCPOLP 9589 bp RNA VRL 20-JUN-1998
DEFINITION Hepatitis C virus genomic RNA for polyprotein, complete cds.
ACCESSION D00944
VERSION D00944.1 GI:221650
KEYWORDS polyprotein.
SOURCE Hepatitis C virus (isolate:HC-J6) cDNA to genomic RNA.
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
REFERENCE 1 (bases 1 to 9589)
AUTHORS Okamoto,H., Okada,S., Sugiyama,Y., Kurai,K., Iizuka,H., Machida,A., Miyakawa,Y. and Mayumi,M.
TITLE Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human carrier: comparison with reported isolates for conserved and divergent regions
JOURNAL J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
MEDLINE 92044440
REFERENCE 2 (sites)
AUTHORS Han,J.H. and Houghton,M.
TITLE Group specific sequences and conserved secondary structures at the 3' end of HCV genome and its implication for viral replication
JOURNAL Nucleic Acids Res. 20 (13), 3520 (1992)
MEDLINE 92335016
REFERENCE 3 (sites)
AUTHORS Hotta,H., Doi,H., Hayashi,T., Purwanta,M., Soemarto,W., Mizokami,M., Ohba,K. and Homma,M.
TITLE Analysis of the core and E1 envelope region sequences of a novel variant of hepatitis C virus obtained in Indonesia
JOURNAL Arch. Virol. 136 (1-2), 53-62 (1994)
MEDLINE 94270990
COMMENT These data kindly submitted in computer readable form by: Hiroaki Okamoto
Immunology Division
Jichi Medical School
Kawachi-gun
Tochigi-ken 329-04
Japan
Phone: 0285-44-2111 x3334
Fax: 0285-44-1557.
FEATURES Location/Qualifiers
source 1. .9589
/organism="Hepatitis C virus"
/isolate="HC-J6"
/db_xref="taxon:11103"
CDS 341. .9442

```
            *    20027    20126: gap of unknown length
            *    20127    22336: contig of 2210 bp in length
            *    22337    22436: gap of unknown length
            *    22437    24423: contig of 1987 bp in length
            *    24424    24523: gap of unknown length
            *    24524    25992: contig of 1469 bp in length
            *    25993    26092: gap of unknown length
            *    26093    28201: contig of 2109 bp in length
            *    28202    28301: gap of unknown length
            *    28302    29954: contig of 1653 bp in length
            *    29955    30054: gap of unknown length
            *    30055    31259: contig of 1205 bp in length
            *    31260    31359: gap of unknown length
            *    31360    33236: contig of 1877 bp in length
            *    33237    33336: gap of unknown length
            *    33337    36092: contig of 2756 bp in length
            *    36093    36192: gap of unknown length
            *    36193    39292: contig of 3100 bp in length
            *    39293    39392: gap of unknown length
            *    39393    41807: contig of 2415 bp in length
            *    41808    41907: gap of unknown length
            *    41908    43021: contig of 1114 bp in length
            *    43022    43121: gap of unknown length
            *    43122    46643: contig of 3522 bp in length
            *    46644    46743: gap of unknown length
            *    46744    49961: contig of 3218 bp in length
            *    49962    50061: gap of unknown length
            *    50062    53523: contig of 3462 bp in length
            *    53524    53623: gap of unknown length
            *    53624    60457: contig of 6834 bp in length
            *    60458    60557: gap of unknown length
            *    60558    64477: contig of 3920 bp in length
            *    64478    64577: gap of unknown length
            *    64578    70891: contig of 6314 bp in length
            *    70892    70991: gap of unknown length
            *    70992    72966: contig of 1975 bp in length
            *    72967    73066: gap of unknown length
            *    73067    81422: contig of 8356 bp in length
            *    81423    81522: gap of unknown length
            *    81523    88228: contig of 6706 bp in length
            *    88229    88328: gap of unknown length
            *    88329    91707: contig of 3379 bp in length
            *    91708    91807: gap of unknown length
            *    91808   100081: contig of 8274 bp in length
            *   100082   100181: gap of unknown length
            *   100182   112695: contig of 12514 bp in length
            *   112696   112795: gap of unknown length
            *   112796   124389: contig of 11594 bp in length
            *   124390   124489: gap of unknown length
            *   124490   145241: contig of 20752 bp in length
            *   145242   145341: gap of unknown length
            *   145342   171394: contig of 26053 bp in length
            *   171395   171494: gap of unknown length
            *   171495   214139: contig of 42645 bp in length.
FEATURES             Location/Qualifiers
     source          1. .214139
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2183B18"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    58620 a  46379 c  46197 g  58900 t   4043 others
ORIGIN


  Query Match           94.1%;  Score 102.6;  DB 2;  Length 214139;
  Best Local Similarity  96.3%;  Pred. No. 0.0037;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 171576 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 171635

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 171636 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 171684


RESULT  14
A08916
LOCUS       A08916         161 bp    DNA             PAT       02-SEP-1993
DEFINITION  H.sapiens (haplotype 3, allele MS32, isolate English, serial number
            32) minisatellite sequence.
ACCESSION   A08916
VERSION     A08916.1  GI:411838
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 99 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .161
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      147 a      0 c      0 g     14 t
ORIGIN


  Query Match           93.8%;  Score 102.2;  DB 6;  Length 161;
  Best Local Similarity  97.2%;  Pred. No. 0.013;
  Matches  104;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      4 AAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     64 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 110


RESULT  15
CNS06EJ3
LOCUS       CNS06EJ3       353 bp    DNA             STS       10-JAN-2001
DEFINITION  T7 end of clone AR0AA022A02 of library AR0AA from strain CBS 732 of
            Zygosaccharomyces rouxii, sequence tagged site.
ACCESSION   AL395173
VERSION     AL395173.1  GI:12146610
KEYWORDS    STS.
SOURCE      Zygosaccharomyces rouxii.
  ORGANISM  Zygosaccharomyces rouxii
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
REFERENCE   1  (bases 1 to 353)
  AUTHORS   de Montigny,J., Straub,M., Potier,S., Tekaia,F., Dujon,B.,
            Wincker,P., Artiguenave,F. and Souciet,J.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 8.
            Zygosaccharomyces rouxii(1)
  JOURNAL   FEBS Lett. 487 (1), 52-55 (2000)
   PUBMED   11152883
REFERENCE   2  (bases 1 to 353)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
            yeast species for molecular evolution studies(1)
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   PUBMED   11152876
```

```
Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||  |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480


RESULT  10
E07362/c
LOCUS       E07362       9589 bp    RNA              PAT       29-SEP-1997
DEFINITION  cDNA of Hepatitis non-A non-B virus.
ACCESSION   E07362
VERSION     E07362.1  GI:2175501
KEYWORDS    JP 1994121689-A/2.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 9589)
  AUTHORS   Okamoto,H. and Nakamura,T. .
  TITLE     NON-A NON-B HEPATITIC VIRUS GENE, POLYNUCLEOTIDE, POLYPEPTIDE,
            ANTIGEN AND ANTIBODY DETECTION SYSTEM
  JOURNAL   Patent: JP 1994121689-A 2 06-MAY-1994;
            NAKAMURA TETSUO
COMMENT     OS   Hepatitis non-A non-B virus
            PN   JP 1994121689-A/2
            PD   06-MAY-1994
            PF   09-AUG-1991 JP 1991287402
            PI   OKAMOTO HIROAKI, NAKAMURA TETSUO
            PC   C12N15/51,A61K39/29,A61K39/395,A61K39/395,C07K13/00,C12P21/02,
            PC   C12P21/08,
            PC   G01N33/53,G01N33/576,G01N33/577//A61B10/00;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .9589
            FT                   /organism='Hepatitis non-A non-B virus' FT
                                 /strain='HCV-J6'.
FEATURES             Location/Qualifiers
     source          1. .9589
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT     1968 a    2820 c    2635 g    2166 t
ORIGIN


  Query Match             94.1%;  Score 102.6;  DB 6;  Length 9589;
  Best Local Similarity   96.3%;  Pred. No. 0.0061;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||  |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480


RESULT  11
I12861/c
LOCUS       I12861       9589 bp    DNA              PAT       26-JUL-1995
DEFINITION  Sequence 2 from patent US 5428145.
ACCESSION   I12861
VERSION     I12861.1  GI:910242
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 9589)
  AUTHORS   Okamoto,H. and Nakamura,T.
  TITLE     Non-A, non-B, hepatitis virus genome, polynucleotides,
            polypeptides, antigen, antibody and detection systems
  JOURNAL   Patent: US 5428145-A 2 27-JUN-1995;
FEATURES             Location/Qualifiers
     source          1. .9589
                     /organism="unknown"
BASE COUNT     1968 a    2820 c    2635 g    2166 t
ORIGIN


  Query Match             94.1%;  Score 102.6;  DB 6;  Length 9589;
  Best Local Similarity   96.3%;  Pred. No. 0.0061;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          ||||  |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

Qy     61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480


RESULT  12
HPCPOLP/c
LOCUS       HPCPOLP      9589 bp    RNA              VRL       20-JUN-1998
DEFINITION  Hepatitis C virus genomic RNA for polyprotein, complete cds.
ACCESSION   D00944
VERSION     D00944.1  GI:221650
KEYWORDS    polyprotein.
SOURCE      Hepatitis C virus (isolate:HC-J6) cDNA to genomic RNA.
  ORGANISM  Hepatitis C virus
            Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
            Hepacivirus.
REFERENCE   1  (bases 1 to 9589)
  AUTHORS   Okamoto,H., Okada,S., Sugiyama,Y., Kurai,K., Iizuka,H., Machida,A.,
            Miyakawa,Y. and Mayumi,M.
  TITLE     Nucleotide sequence of the genomic RNA of hepatitis C virus
            isolated from a human carrier: comparison with reported isolates
            for conserved and divergent regions
  JOURNAL   J. Gen. Virol. 72 (Pt 11), 2697-2704 (1991)
  MEDLINE   92044440
REFERENCE   2  (sites)
  AUTHORS   Han,J.H. and Houghton,M.
  TITLE     Group specific sequences and conserved secondary structures at the
            3' end of HCV genome and its implication for viral replication
  JOURNAL   Nucleic Acids Res. 20 (13), 3520 (1992)
  MEDLINE   92335016
REFERENCE   3  (sites)
  AUTHORS   Hotta,H., Doi,H., Hayashi,T., Purwanta,M., Soemarto,W.,
            Mizokami,M., Ohba,K. and Homma,M.
  TITLE     Analysis of the core and E1 envelope region sequences of a novel
            variant of hepatitis C virus obtained in Indonesia
  JOURNAL   Arch. Virol. 136 (1-2), 53-62 (1994)
  MEDLINE   94270990
COMMENT     These data kindly submitted in computer readable form by:  Hiroaki
            Okamoto
            Immunology Division
            Jichi Medical School
            Kawachi-gun
            Tochigi-ken 329-04
            Japan
            Phone:  0285-44-2111 x3334
            Fax:    0285-44-1557.
FEATURES             Location/Qualifiers
     source          1. .9589
                     /organism="Hepatitis C virus"
                     /isolate="HC-J6"
                     /db_xref="taxon:11103"
     CDS             341. .9442
```

```
                     /codon_start=1
                     /product="polyprotein"
                     /protein_id="BAA00792.1"
                     /db_xref="GI:221651"
                     /translation="MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRL
                     GVRATRKTSERSQPRGRRQPIPKDRRSTGKSWGKPGYPWPLYGNEGLGWAGWLLSPRG
                     SRPSWGPNDPRHRSRNVGKVIDTLTCGFADLMGYIPVVGAPLGGVARALAHGVRVLED
                     GVNFATGNLPGCSFSIFLLALLSCITTPVSAAEVKNISTGYMVTNDCTNDSITWQLQA
                     AVLHVPGCVPCEKVGNTSRCWIPVSPNVAVQQPGALTQGLRTHIDMVVMSATLCSALY
                     VGDLCGGVMLAAQMFIVSPQHHWFVQDCNCSIYPGTITGHRMAWDMMMNWSPTATMIL
                     AYAMRVPEVIIDIIGGAHWGVMFGLAYFSMQGAWAKVVVILLLAAGVDAQTHTVGGST
                     AHNARTLTGMFSLGARQKIQLINTNGSWHINRTALNCNDSLHTGFLASLFYTHSFNSS
                     GCPERMSACRSIEAFRVGWGALQYEDNVTNPEDMRPYCWHYPPRQCGVVSASSVCGPV
                     YCFTPSPVVVGTTDRLGAPTYTWGENETDVFLLNSTRPPQGSWFGCTWMNSTGYTKTC
                     GAPPCRIRADFNASMDLLCPTDCFRKHPDTTYIKCGSGPWLTPRCLIDYPYRLWHYPC
                     TVNYTIFKIRMYVGGVEHRLTAACNFTRGDRCNLEDRDRSQLSPLLHSTTEWAILPCT
                     YSDLPALSTGLLHLHQNIVDVQFMYGLSPALTKYIVRWEWVVLLFLLLADARVCACLW
                     MLILLGQAEAALEKLVVLHAASAASCNGFLYFVIFFVAAWYIKGRVVPLATYSLTGLW
                     SFGLLLLALPQQAYAYDASVHGQIGAALLVLITLFTLTPGYKTLLSRFLWWLCYLLTL
                     AEAMVQEWAPPMQVRGGRDGIIWAVAIFCPGVVFDITKWLLAVLGPAYLLKGALTRVP
                     YFVRAHALLRMCTMVRHLAGGRYVQMVLLALGRWTGTYIYDHLTPMSDWAANGLRDLA
                     VAVEPIIFSPMEKKVIVWGAETAACGDILHGLPVSARLGREVLLGPADGYTSKGWSLL
                     APITAYAQQTRGLLGTIVVSMTGRDKTEQAGEIQVLSTVTQSFLGTTISGVLWTVYHG
                     AGNKTLAGSRGPVTQMYSSAEGDLVGWPSPPGTKSLEPCTCGAVDLYLVTRNADVIPA
                     RRRGDKRGALLSPRPLSTLKGSSGGPVLCPRGHAVGVFRAAVCSRGVAKSIDFIPVET
                     LDIVTRSPTFSDNSTPPAVPQTYQVGYLHAPTGSGKSTKVPVAYAAQGYKVLVLNPSV
                     AATLGFGAYLSKAHGINPNIRTGVRTVTTGAPITYSTYGKFLADGGCAGGAYDIIICD
                     ECHAVDSTTILGIGTVLDQAETAGVRLTVLATATPPGSVTTPHPNIEEVALGQEGEIP
                     FYGRAIPLSYIKGGRHLIFCHSKKKCDELAAALRGMGLNAVAYYRGLDVSVIPTQGDV
                     VVVATDALMTGFTGDFDSVIDCNVAVTQVVDFSLDPTFTITTQTVPQDAVSRSQRRGR
                     TGRGRLGIYRYVSTGERASGMFDSVVLCECYDAGAAWYELTPAETTVRLRAYFNTPGL
                     PVCQDHLEFWEAVFTGLTHIDAHFLSQTKQSGENFAYLTAYQATVCARAKAPPPSWDV
                     MWKCLTRLKPTLVGPTPLLYRLGSVTNEVTLTHPVTKYIATCMQADLEVMTSTWVLAG
                     GVLAAVAAYCLATGCVCIIGRLHVNQRAVVAPDKEVLYEAFDEMEECASRAALIEEGQ
                     RIAEMLKSKIQGLLQQASKQAQDIQPAVQASWPKVEQFWAKHMWNFISGIQYLAGLST
                     LPGNPAVASMMAFSAALTSPLSTSTTILLNILGGWLASQIAPPAGATGFVVSGLVGAA
                     VGSIGLGKVLVDILAGYGAGISGALVAFKIMSGEKPSMEDVVNLLPGILSPGALVVGV
                     ICAAILRRHVGPGEGAVQWMNRLIAFASRGNHVAPTHYVTESDASQRVTQLLGSLTIT
                     SLLRRLHNWITEDCPIPCSGSWLRDVWDWVCTILTDFKNWLTSKLFPKMPGLPFISCQ
                     KGYKGVWAGTGIMTTRCPCGANISGNVRLGSMRITGPKTCMNIWQGTFPINCYTEGQC
                     VPKPAPNFKIAIWRVAASEYAEVTQHGSYHYITGLTTDNLKVPCQLPSPEFFSWVDGV
                     QIHRFAPIPKPFFRDEVSFCVGLNSFVVGSQLPCDPEPDTDVLTSMLTDPSHITAETA
                     ARRLARGSPPSEASSSASQLSAPSLRATCTTHGKAYDVDMVDANLFMGGDVTRIESES
                     KVVVLDSLDPMVEERSDLEPSIPSEYMLPKKRFPPALPAWARPDYNPPLVESWKRPDY
                     QPATVAGCALPPPKKTPTPPPRRRRTVGLSESSIADALQQLAIKSFGQPPPSGDSGLS
                     TGADAADSGSRTPPDELALSETGSISSMPPLEGEPGDPDLEPEQVELQPPPQGGVVTP
                     GSGSGSWSTCSEEDDSVVCCSMSYSWTGALITPCSPEEEKLPINPLSNSLLRYHNKVY
                     CTTSKSASLRAKKVTFDRMQALDAHYDSVLKDIKLAASKVTARLLTLEEACQLTPPHS
                     ARSKYGFGAKEVRSLSGRAVNHIKSVWKDLLEDTQTPIPTTIMAKNEVFCVDPTKGGK
                     KAARLIVYPDLGVRVCEKMALYDITQKLPQAVMGASYGFQYSPAQRVEFLLKAWAEKK
                     DPMGFSYDTRCFDSTVTERDIRTEESIYRACSLPEEAHTAIHSLTERLYVGGPMFNSK
                     GQTCGYRRCRASGVLTTSMGNTITCYVKALAACKAAGIIAPTMLVCGDDLVVISESQG
                     TEEDERNLRAFTEAMTRYSAPPGDPPRPEYDLELITSCSSNVSVALGPQGRRRYYLTR
                     DPTTPIARAAWETVRHSPVNSWLGNIIQYAPTIWARMVLMTHFFSILMAQDTLDQNLN
                     FEMYGAVYSVSPLDLPAIIERLHGLDAFSLHTYTPHELTRVASALRKLGAPPLRAWKS
                     RARAVRASLISRGGRAAVCGRYLFNWAVKTKLKLTPLPEARLLDLSSWFTVGAGGGDI
                     YHSVSRARPRLLLLGLLLFVGVGLFLLPAR"
ASE COUNT      1970 a    2820 c    2633 g    2166 t
RIGIN


 Query Match             94.1%;  Score 102.6;  DB 14;  Length 9589;
 Best Local Similarity   96.3%;  Pred. No. 0.0061;
 Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

y       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
          |||| |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||
o    9588 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 9529

y      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
b    9528 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 9480
```

```
RESULT  13
AC022129
LOCUS       AC022129   214139 bp    DNA              HTG       20-APR-2001
DEFINITION  Homo sapiens chromosome 5 clone CTD-2183B18, WORKING DRAFT
            SEQUENCE, 40 unordered pieces.
ACCESSION   AC022129
VERSION     AC022129.5  GI:13699626
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 214139)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 214139)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (26-JAN-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 20, 2001 this sequence version replaced gi:9954660.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 697863
            Center clone name: CITB-H1_2183B18
            -------
            Summary Statistics
            Consensus quality: 177232 bases at least Q40
            Consensus quality: 190777 bases at least Q30
            Consensus quality: 195469 bases at least Q20
            Estimated insert size: 121000; pulse field gel estimation
            Estimated insert size: 210239; sum-of-contigs estimation
            Quality coverage:  7.99 in Q20 bases; pulse field gel estimation
            Quality coverage:  4.6 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 40 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1159: contig of 1159 bp in length
            *     1160     1259: gap of unknown length
            *     1260     2521: contig of 1262 bp in length
            *     2522     2621: gap of unknown length
            *     2622     3853: contig of 1232 bp in length
            *     3854     3953: gap of unknown length
            *     3954     5111: contig of 1158 bp in length
            *     5112     5211: gap of unknown length
            *     5212     6367: contig of 1156 bp in length
            *     6368     6467: gap of unknown length
            *     6468     7531: contig of 1064 bp in length
            *     7532     7631: gap of unknown length
            *     7632     8836: contig of 1205 bp in length
            *     8837     8936: gap of unknown length
            *     8937    10072: contig of 1136 bp in length
            *    10073    10172: gap of unknown length
            *    10173    11593: contig of 1421 bp in length
            *    11594    11693: gap of unknown length
            *    11694    13594: contig of 1901 bp in length
            *    13595    13694: gap of unknown length
            *    13695    15802: contig of 2108 bp in length
            *    15803    15902: gap of unknown length
            *    15903    17998: contig of 2096 bp in length
            *    17999    18098: gap of unknown length
            *    18099    20026: contig of 1928 bp in length
```

```
            *    20027    20126: gap of unknown length
            *    20127    22336: contig of 2210 bp in length
            *    22337    22436: gap of unknown length
            *    22437    24423: contig of 1987 bp in length
            *    24424    24523: gap of unknown length
            *    24524    25992: contig of 1469 bp in length
            *    25993    26092: gap of unknown length
            *    26093    28201: contig of 2109 bp in length
            *    28202    28301: gap of unknown length
            *    28302    29954: contig of 1653 bp in length
            *    29955    30054: gap of unknown length
            *    30055    31259: contig of 1205 bp in length
            *    31260    31359: gap of unknown length
            *    31360    33236: contig of 1877 bp in length
            *    33237    33336: gap of unknown length
            *    33337    36092: contig of 2756 bp in length
            *    36093    36192: gap of unknown length
            *    36193    39292: contig of 3100 bp in length
            *    39293    39392: gap of unknown length
            *    39393    41807: contig of 2415 bp in length
            *    41808    41907: gap of unknown length
            *    41908    43021: contig of 1114 bp in length
            *    43022    43121: gap of unknown length
            *    43122    46643: contig of 3522 bp in length
            *    46644    46743: gap of unknown length
            *    46744    49961: contig of 3218 bp in length
            *    49962    50061: gap of unknown length
            *    50062    53523: contig of 3462 bp in length
            *    53524    53623: gap of unknown length
            *    53624    60457: contig of 6834 bp in length
            *    60458    60557: gap of unknown length
            *    60558    64477: contig of 3920 bp in length
            *    64478    64577: gap of unknown length
            *    64578    70891: contig of 6314 bp in length
            *    70892    70991: gap of unknown length
            *    70992    72966: contig of 1975 bp in length
            *    72967    73066: gap of unknown length
            *    73067    81422: contig of 8356 bp in length
            *    81423    81522: gap of unknown length
            *    81523    88228: contig of 6706 bp in length
            *    88229    88328: gap of unknown length
            *    88329    91707: contig of 3379 bp in length
            *    91708    91807: gap of unknown length
            *    91808   100081: contig of 8274 bp in length
            *   100082   100181: gap of unknown length
            *   100182   112695: contig of 12514 bp in length
            *   112696   112795: gap of unknown length
            *   112796   124389: contig of 11594 bp in length
            *   124390   124489: gap of unknown length
            *   124490   145241: contig of 20752 bp in length
            *   145242   145341: gap of unknown length
            *   145342   171394: contig of 26053 bp in length
            *   171395   171494: gap of unknown length
            *   171495   214139: contig of 42645 bp in length.
FEATURES             Location/Qualifiers
     source          1. .214139
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTD-2183B18"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    58620 a  46379 c  46197 g  58900 t   4043 others
ORIGIN


  Query Match             94.1%;  Score 102.6;  DB 2;  Length 214139;
  Best Local Similarity   96.3%;  Pred. No. 0.0037;
  Matches  105;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| ||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db  171576 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 171635

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaagg 109
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  171636 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGG 171684


RESULT  14
A08916
LOCUS       A08916        161 bp    DNA             PAT       02-SEP-1993
DEFINITION  H.sapiens (haplotype 3, allele MS32, isolate English, serial number
            32) minisatellite sequence.
ACCESSION   A08916
VERSION     A08916.1  GI:411838
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161)
  AUTHORS   Jeffreys,A.J.
  TITLE     Extended nucleotide sequences
  JOURNAL   Patent: EP 0370719-A 99 30-MAY-1990;
            IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES             Location/Qualifiers
     source          1. .161
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      147 a      0 c      0 g     14 t
ORIGIN


  Query Match             93.8%;  Score 102.2;  DB 6;  Length 161;
  Best Local Similarity   97.2%;  Pred. No. 0.013;
  Matches  104;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           ||||||||||   ||||||||||||||||||||||||||||||||||||||||||||||
Db       4 AAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 63

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db      64 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 110


RESULT  15
CNS06EJ3
LOCUS       CNS06EJ3      353 bp    DNA             STS       10-JAN-2001
DEFINITION  T7 end of clone AR0AA022A02 of library AR0AA from strain CBS 732 of
            Zygosaccharomyces rouxii, sequence tagged site.
ACCESSION   AL395173
VERSION     AL395173.1  GI:12146610
KEYWORDS    STS.
SOURCE      Zygosaccharomyces rouxii.
  ORGANISM  Zygosaccharomyces rouxii
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Zygosaccharomyces.
REFERENCE   1  (bases 1 to 353)
  AUTHORS   de Montigny,J., Straub,M., Potier,S., Tekaia,F., Dujon,B.,
            Wincker,P., Artiguenave,F. and Souciet,J.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 8.
            Zygosaccharomyces rouxii(1)
  JOURNAL   FEBS Lett. 487 (1), 52-55 (2000)
   PUBMED   11152883
REFERENCE   2  (bases 1 to 353)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
            yeast species for molecular evolution studies(1)
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
   PUBMED   11152876
```

```
REFERENCE   3  (bases 1 to 353)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This STS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .353
                     /organism="Zygosaccharomyces rouxii"
                     /strain="CBS 732"
                     /db_xref="taxon:4956"
                     /clone="AR0AA022A02"
                     /clone_lib="AR0AA"
                     /note="end : T7"
     STS             1. .353
BASE COUNT      314 a     12 c     18 g      7 t      2 others
ORIGIN


     Match             93.8%;  Score 102.2;  DB 11;  Length 353;
     Local Similarity  97.2%;  Pred. No. 0.011;
       s  104;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 aaaataaaaatttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 60
           |||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       5 AAAAAAAGAAATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 304

Qy      61 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 107
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     305 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 351


Search completed: January  3, 2002, 09:21:50
Job time: 1493 sec
```